

Please scan both sides

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame plus n2p model

Run on: September 10, 2004, 14:19:44 ; Search time 68.5 Seconds  
(without alignments)  
7473.450 Million cell updates/sec

```

Title:          VS-10-039-272-1
Perfect score:  4850
Sequence:       1 CGGCAGAGGGCCAGGGA.....aaaaaaaaaaaaaaaaaaaa 2661

```

Scoring table: BLOSUM62

Xgapop	10.0	Xgapext	0.5
Ygapop	10.0	Ygapext	0.5
Fgapop	6.0	Fgapext	7.0
Delop	6.0	Delext	7.0

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 566732

Minimum DB seq length: 0

maxlen DV seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```

MODEL=frames+ n2p. model /US10039272/runat_07092004_144223_21649/app_query.fasta_1.2823
Q=/cgn2_1/UsPto_spool /US10039272
DB=PIR_78 -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
UNITS=bits -START=1 -END=1 -MATRIX=Blosum62 -TRANS=human40.cdi -LIST=45
LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
OUTFMT=pt0 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
USER=US10039272_@CEN 1 1 90 @runat_07092004_144223_21649 -NCPUS=6 -ICPU=3
NO MMAP -LARGEOBJS -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=0.5 -XGAPEXT=7
XGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

```
Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

result No.	Score	% Match		Length	DB	ID	Description
1	2999	61.8		560	2	I38065	Gene NMB protein
2	635	13.1		626	2	S53871	Pmel 17 protein -
3	593.5	12.2		662	2	I38400	melanoma-associate
4	588.5	12.1		668	2	A41234	melanocyte-specific
5	436.5	9.0		491	2	A49179	melanoma antigen h
6	135	2.8		446	2	T07907	hydroxyproline-ric
7	129	2.7		926	1	A41105	protein-tyrosine-p
8	118	2.4		2869	2	T18518	apolipoprotein(a) 9
9	116.5	2.4		555	2	S21766	dihydrolipoamide S
10	114	2.4		2395	1	S08290	surface protein ty
11	113	2.3		1541	2	T02831	AAA protein L4171.
12	112.5	2.3		457	2	I55976	dihydrolipoamide S
13	111.5	2.3		2946	2	T15840	hypothetical prote
14	111	2.3		525	1	KGHUGH	histidine-rich gly

15	110.5	2.3	1874	1	J00533	genome polyprotein
16	109.5	2.3	492	2	C96521	protein F21D18.18
17	109	2.2	1737	2	A59235	unconventional myo
18	108.5	2.2	348	2	AB3260	hypothetical membr
19	108.5	2.2	588	2	T45564	hypothetical prote
20	107.5	2.2	768	2	A87722	protein zc123.1 [1
21	107.5	2.2	1208	2	T27822	hypothetical prote
22	106	2.2	658	2	T08133	hypothetical prote
23	106	2.2	1537	2	S53465	cysteine proteinas
24	106	2.2	1585	2	T31611	floculation prote
25	105	2.2	499	2	S52432	hypothetical prote
26	104.5	2.2	4006	2	T09070	chitinase (EC 3.2.
27	104	2.1	435	2	D41602	probable tenascin
28	103	2.1	626	1	NBHU1A	transcript fact
29	103	2.1	921	2	A33718	platelet glycoprot
30	103	2.1	1009	2	G64483	retinoblastoma pro
31	102.5	2.1	394	2	E82572	hypothetical prote
32	102.5	2.1	826	2	G90283	ABC transporter so
33	102.5	2.1	979	2	A35913	hypothetical prote
34	102	2.1	281	2	T29150	regulatory factor
35	102	2.1	393	2	T33103	hypothetical prote
36	102	2.1	1213	2	A41124	lin-1 protein - Ca
37	102	2.1	2014	2	T21560	limb deformity (ld
38	101.5	2.1	445	2	D81716	hypothetical prote
39	101.5	2.1	2165	2	T21371	hypothetical prote
40	101	2.1	848	2	T23694	hypothetical prote
41	101	2.1	2090	2	S26058	probable transform
42	100.5	2.1	221	2	T07176	extensin homolog -
43	100.5	2.1	1048	2	T30815	platelet-derived g
44	100	2.1	432	2	AB2222	twitching motility
45	100	2.1	26926	1	I38344	titin cardiac mus

## ALIGNMENTS

## RESULT 1

I38065  
 gene NMB protein - human  
 CSpecies: Homo sapiens (man)  
 CDate: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 05-Nov-1999  
 CAccession: I38065  
 R.Weterman, M.A.; Ajubi, N.; van Dinter, I.M.; Degen, W.G.; van Muijen, G.N.; Ruitter, I.  
 Int. J. Cancer 60, 73-81, 1995  
 A>Title: nmb, a novel Gene, is expressed in low-metastatic human melanoma cell lines and  
 A|Reference number: I38065; MUID:95113576; PMID:7814155  
 A|Accession: I38065  
 A|Status: preliminary; translated from GB/EMBL/DBJ  
 A|Molecule type: mRNA  
 A|Residues: 1-360 <RES>  
 A|Cross-references: EMBL:X75334; NID:G666042; PIDN:CAA54044.1; PID:G666043  
 C|Genetics:  
 A|Gene: GDB:NMB  
 A|Cross-references: GDB:I20237; OMIM:162340  
 A|Map position: 15q22-15qter

```

Alignment Scores:
  Aligned No.:      9 66-252
  Score:            2995.00
  Percent Similarity: 99.82%
  Best Local Similarity: 99.82%
  Query Match:      91.84%
  DB:                2
  Length:            560
  Matches:           559
  Conservative:      0
  Mismatches:        1
  Indels:            1
  Gaps:              0

```

US-10-039-272-1 (1-2661) x I38065 (1-560)

60	ATGGAAATGCTCTACATATTCTCTGGGATTTCTGCTCTGGCTGCAAGATTGCCACTTGAT	119
1	MeGluCysLeuTyrTyrPheLeuGlyPheLeuLeuLeuAalaArgLeuProLeuAsp	20
120	GCSCCAACGATTTTCATGCTGCTGGGCAATGAAGACCTTCTGCTTACATGAGGAG	179
21	AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu	40

```
QY 180 CACAAATCAATTAATGCTGCTCTTCATGAAATGACATGAATGAAATCTTACCA 239
|||
Db 41 HisAsnGlnLeuAsnGlyTrpSerAspGluAsnAspTrpAsnGluLeuTyPro 60
|||
QY 240 GTGTGGAAAGCGGGAGACATCAGAGTGGAAATCTCTGGAAAGGAGGCGCGTGTGAGCG 299
|||
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla 80
|||
QY 300 GTCCTGACAGTACTACAGCCCTCTGCGGCTCAATATAACATTTGCGGTGAACCTG 359
|||
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
|||
QY 360 ATATTCCTAGATGCCAAAGAGATGCCAATGCCAATAGTCTATGAGAAGAACTGC 419
|||
Db 101 IlePheProArgGlyGlnLeuGluAspAlaAsnGlyAsnIleValTyGluLysAsnCys 120
|||
QY 420 AGAATAGGCTGGTGTATCTGCTCATCATATGTTTACATGTCAGACAGCATGTCAGAG 479
|||
Db 121 ArgAsnGlnAlaGlyLeuSerAlaAspProTyValTyAsnTrpThrAlaTrpSerGlu 140
|||
QY 480 GACAGTACGGGAAATGGACCGGCCAAAGCCATCATAGCTCTCTGCTGATGGAA 539
|||
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
|||
QY 540 CTTTTCCTCACCACCCGGATGGAGATGGAATTCATCTAGCTCTTCCACACACTT 599
|||
Db 161 ProPheProHisProGlyTrpArgArgTrpAsnPheIleTyValPheHisThrLeu 180
|||
QY 600 GGTCAAGTATTCAGAAATGGAGCATGTTTCTGAGTGGAGATTTCTGTGAACACAGCCAA 659
|||
Db 181 GlyGlnTyPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
|||
QY 660 GTGACACTGGGCTCAACTCATGGAAGTCACTGTCTACAGAGACATGAGCGGCATAT 719
|||
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyArgArgHisGlyArgAlaTy 220
|||
QY 720 GTTCCCATCGCAAGTGAAGATGTAGCTGGTAAACAGATCAGATTCCTGTGTTGTG 779
|||
Db 221 ValProIleAlaGlnValLysAspValTyValThrAspGlnIleProValPheVal 240
|||
QY 780 ACTATCTTCCAGAAAGACCATCAAAATTCATCCGAGCAACCTTGCCTCAAGATCTCC 838
|||
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
|||
QY 839 ATTATGTTGATCTCTGATTCATGATCTAGCATCTTCTGATCTTCTTCTTCTTCTTCT 898
|||
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTySerThrIleAsn 280
|||
QY 899 TACAACTGGAGCTTCGGGATATATCTGGCTGTTGTTTCCACCAATCATCTGTGAAT 958
|||
Db 281 TyTrpLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
|||
QY 959 CACACGTATGTCTCAATGAAACCTTCACTCAGCTTAACTCCTCACTGTAAGCTGAGACCA 1018
|||
Db 301 HisThrTyValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
|||
QY 1019 GGAACCTTGTCCGACCGCCACACACAGACCTTCAAAACCCACCCCTCTTTAGGA 1078
|||
Db 321 GlyProCysProProProProProProProProProProProProProProProPro 340
|||
QY 1079 CTTGCTGTCACAAACCCCTGGAGCTGAGTAGGATTCCTCATGAAATCTCCAGATTAAC 1138
|||
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
|||
QY 1139 AGATATGGCCACTTCAAGCCACCATCACAATTTAGAGGAATTTTAGAGGTTTAAATC 1198
|||
Db 361 ArgTyTrpGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
|||
QY 1199 ATCCAGATGACAGACGCTCTGATCCCGTCCGCTGAAAGCTCCCTAAATAGACTTT 1258
|||
Db 381 IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe 400
|||
QY 1259 GTCTGTGACCTGCCAAGGAGCATTCACACGAGGCTGTGTACCATCATTTCTGACCCACC 1318
|||
```

## RESULT 2

S53871

Pmel 17 protein - mouse

C:Species: Mus musculus (house mouse)

C:Date: 27-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 05-Nov-1999

C:Accession: S53871

R:Kwon, B.S.; Halaban, R.; Ponnazhagan, S.; Kim, K.; Chintamaneni, C.; Bennett, D.;

Nucleic Acids Res. 23, 154-158, 1995

A:Title: Mouse silver mutation is caused by a single base insertion in the putative cy

A:Reference number: S53871; MUID:95175358; PMID:7870580

A:Accession: S53871

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-626 &lt;KWO&gt;

A:Cross-references: GB:U14133; NID:5887940; PIDN:AAA69538.1; PID:9887941

Alignment Scores:

Pred. No.: 1.49e-46 Length: 626

Score: 635.00 Matches: 184

Percent Similarity: 41.45% Conservative: 97

Best Local Similarity: 27.14% Mismatches: 207

Query Match: 13.09% Indels: 191

DB: 2 Gaps: 22

US-10-039-272-1 (1-2661) x S53871 (1-626)

```
QY 78 TTCCTGGATTTCTGCTCTGCTGCTGCAAGATTGCCACTTGATGCC-----GCCAAACGA 131
|||
Db 9 PheLeuProValLeuValLeuSerAlaLeuLeuAlaValGlyAlaLeuGluGlySerArg 28
|||
QY 132 TTCATCATGTGCTGGCGCAATGAAGACCTCTGCTTACATGAGGAGGACCAATCAATTA 191
|||
Db 29 AsnGlnAspTrpLeuGlyValProArgGlnLeuVal----- 40
|||
QY 192 AATGGCTGGTCTTCTGATGAAATGACTGGAATGAAAACTTACCCAGTGTGAACGG 251
|||
Db 41 -----ThrLysThrTrpAsnArgGlnLeuTyProGluTrpThr--- 53
|||
QY 252 GGAGACATGAGTGGAAAAAATCTCTGGAAGGAGGCGCTGTGAGGCGGCTCTGACAGT 311
|||
Db 54 ---GluValGlnGlySerAsnCysTrpArgGlyGlnValSerLeuArgValIleAsn 72
|||
```

1260 TGGCATGGCTGAAGCTCCCTAATAGACTTTGTGCTGACCTGCCAAGGAGCATTTCCCA 1319  
1287 CGGAGTCTGTACCATCTTCTGACCCCACTGCGAGATCACCCAGAACACAGTCTGCA 1346  
1320 CGGAGTCTGTACCATCTTCTGACCCCACTGCGAGATCACCCAGAACACAGTCTGCA 1379  
1347 GCGCTGTGATGTGATGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1406  
1380 GCGCTGTGATGTGATGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1439  
1407 GAGCTGTGATGTGATGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1466  
1440 GAGCTGTGATGTGATGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1499  
1467 TGAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1526  
1500 TGAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1559  
1527 TCTCGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1586  
1560 TCTCGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1619  
1587 ACAAGGATACAAACCAATAGAAAATAGTCTGGAATGTGCTGGAAGCAAGGCTGCA 1646  
1620 ACAAGGATACAAACCAATAGAAAATAGTCTGGAATGTGCTGGAAGCAAGGCTGCA 1679  
1647 GTGCTCTTCTCAACCGTCAAAAGCGGTCTTCTCCGGGAACACAGAAAGGATCGCG 1706  
1680 GTGCTCTTCTCAACCGTCAAAAGCGGTCTTCTCCGGGAACACAGAAAGGATCGCG 1739  
1707 TACTCAAAAACCAAGAAATTAAGAGGTTCTTAAATTTGACCTGTTCTGAGCTCA 1766  
1740 TACTCAAAAACCAAGAAATTAAGAGGTTCTTAAATTTGACCTGTTCTGAGCTCA 1799  
1767 CTCTTCAGTGCAATGATGTGAGATGTGCTGGAGTGGCTATTAACTTTTCTTAAAG 1826  
1800 CTCTTCAGTGCAATGATGTGAGATGTGCTGGAGTGGCTATTAACTTTTCTTAAAG 1859  
1827 ATTAATGTTAATAGATTTGTTTGGGGAAGTGAATTTTATAGGTTAAATGCA 1886  
1860 ATTAATGTTAATAGATTTGTTTGGGGAAGTGAATTTTATAGGTTAAATGCA 1919  
1887 TTTTAGATGGGAGAGGATTAATCTGACGAGGCTTCAAGCTGTTGTTGAACTGAT 1946  
1920 TTTTAGATGGGAGAGGATTAATCTGACGAGGCTTCAAGCTGTTGTTGAACTGAT 1979  
1947 AAAAGCACTTAGCAAGCTCTTTTCAATATTTTATGTTTCACTTATAAGTCTTAG 2006  
1980 AAAAGCACTTAGCAAGCTCTTTTCAATATTTTATGTTTCACTTATAAGTCTTAG 2039  
2007 GTAACTAGTAGGATAGAAAACATCTGTCGAGAGTAAAGGAGAGGCTTCACTTAGTA 2066  
2040 GTAACTAGTAGGATAGAAAACATCTGTCGAGAGTAAAGGAGAGGCTTCACTTAGTA 2099  
2067 GAGCTAACCCAGGTTAACTGCAAGAGAGCGGGATCTTTGAGCTTTCACTGTAAGT 2126  
2100 GAGCTAACCCAGGTTAACTGCAAGAGAGCGGGATCTTTGAGCTTTCACTGTAAGT 2159  
2127 TATGCTAAAGCCATGTAGTCCAGTTTCAAGATCAATGTTCCAGCTTAACTGAATCCCA 2186  
2160 TATGCTAAAGCCATGTAGTCCAGTTTCAAGATCAATGTTCCAGCTTAACTGAATCCCA 2219  
2187 CTTCATATACACTCATGAACTCTGATGAGCAATTAACAGGCTTAACTGAATCCCA 2246  
2220 CTTCATATACACTCATGAACTCTGATGAGCAATTAACAGGCTTAACTGAATCCCA 2279  
2247 TGTGCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2306  
2280 TGTGCACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2339  
2307 GTGCAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2366

## RESULT 5

\*  
HSNMB

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

2340 GTGACAACTTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCATTATTCCCA 2399  
2367 TGGACATTTAGTTAGTCTTTTATATACAGGATGCTGCTGAGTGAACACTCTTTGTGA 2426  
2400 TGGACATTTAGTTAGTCTTTTATATACAGGATGCTGCTGAGTGAACACTCTTTGTGA 2459  
2427 TATTTCCAAATTTTGTATAGTCTGCTGACATATTTGAAATCAAAATATTAAGACTTTCC 2486  
2460 TATTTCCAAATTTTGTATAGTCTGCTGACATATTTGAAATCAAAATATTAAGACTTTCC 2518  
2487 AAAAATTTGCTCCCTGTTTTCATGCGAAGTCTGATCAGTAAGGATTTCCCTCTCTTTG 2546  
2519 AAAGATGAGTCCCTGTTTTCATGCGAAGTCTGATCAGTAAGGATTTCCCTCTCTTTG 2578  
2547 GAACTAAACCACTTACTATATTTAGTGAAGCAAGACATTTTCTTCTCTCTCTCTCTCTCT 2606  
2579 TAACTAAACCACTTACTATATTTAGTGAAGCAAGACATTTTCTTCTCTCTCTCTCTCTCT 2638  
2607 AAAATGAGGGAAGACAAAAA 2636  
2639 TAAAGTGTGGGAAGACAAAAA 2669

HSNMB

H.sapiens NMB mRNA.

X76534

X76534.1 GI:666042

NMB gene.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2669)

Weterman, M.A., Ajubi, N., van Dinter, I.M., Degen, W.G., van

Muijen, G.N., Rutter, D.J. and Bloemers, H.P.

nmb, a novel gene, is expressed in low-metastatic human melanoma

cell lines and xenografts

Int. J. Cancer 60 (1), 73-81 (1995)

95113576

7814155

2

Weterman, M.

Direct Submission

Submitted (03-DEC-1993)

M. Weterman, University of Nijmegen, Dept

of Biochemistry, PO Box 9101, 6500 HB Nijmegen, NETHERLANDS

Location/Qualifiers

1. .2669

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="NMB"

/cell\_line="MV1"

/clone\_lib="MV1"

1. .2669

/gene="NMB"

92. .1774

/gene="NMB"

/codon\_start=1

/protein\_id="CAA54044.1"

/db\_xref="GI:666043"

/db\_xref="GOA:Q14956"

/translation="MECLYFLGLLLAARLPDAAKRFHVLGNRPAYRHNQL

NWSDDENNEKLYPVKRGDMRWKNSWQVAVLTSDSPALVGSNITFVNLIF

PRCQKEDANGNIYKNCNREAGLSADPYVYNWTAWSDESGENGSHHNVFDPK

PFPHIPGWRMNFIVFHTLGOYFQKIGRCSVRVSUNTANVTLPGLMVTYRHRGR

AVVPAOVKDVVVTDOIIEVFTMFQKNDNSDETFLKDLFIMEDVLHDPHFLNY

STINKWSGDNTGLFVSTNHNHYVINGTFSNLTVAAPAGCPPPPSPSK

PTSPGPDGNEELSRIPDENCINRYGHQATITVEGILEVNIQMTDVLMPVPM

PSSLLDFVTCQSGIPEVCTIISDPTCEITQNTVCPVDVDEMLCLTVRTFNGSG

TYCVNLTLGDDTSLAUTSLTSLVSPDRDPASPLRNANSALISVGCLAFVTVLSLLVYK







Query Match 94.8%; Score 2522; DB 9; Length 2787; Best Local Similarity 97.8%; Pred. No. 0; Matches 2613; Conservative 0; Mismatches 20; Indels 39; Gaps 4;									
QY	28	AACCTTGGTGGCTGGTCCGTCGAGAAATTCAGCATGGAATGTCTACTATTTCCTGGGAT	87	1047	CCAGACCTTCAAAACCCACCCCTTCTTT	-----	1074		
DB	46	ACCTTGGTGGCTGGTCCGTCGAGAAATTCAGCATGGAATGTCTACTATTTCCTGGGAT	105	1066	CCAGACCTTCAAAAGCCACCCCTTCTTTAGCAACTACTCTAAAACTTATGATTCAAAACA	1125			
QY	88	TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAAACGAATTCATGATGCTGG	147	1075	-----AGACCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCC	1130			
DB	106	TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAAACGAATTCATGATGCTGG	165	1126	CCCAGGACCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCC	1185			
QY	148	GCAATGAAAGACCTTCTGCTTACATGAGGAGACAAATCAATTAATGCTGGTCTTCTG	207	1131	AGATTACAGATATGGCCACTTTCAAGCCACCATCACAAATGTAGAGGGAATCTTAGAGG	1190			
DB	166	GCAATGAAAGACCTTCTGCTTACATGAGGAGACAAATCAATTAATGCTGGTCTTCTG	225	1186	AGATTACAGATATGGCCACTTTCAAGCCACCATCACAAATGTAGAGGGAATCTTAGAGG	1245			
QY	208	ATGAAATGACTGGAAATGAAACCTTACCCAGTGTGGAAGCGGGAGACATGAGGTGA	267	1191	TTAAACATCATCCAGATGACAGACGCTCTGATGCCGGTGCCATGGCCTGAAAGCTCCCTAA	1250			
DB	226	ATGAAATGACTGGAAATGAAACCTTACCCAGTGTGGAAGCGGGAGACATGAGGTGA	285	1246	TTAAACATCATCCAGATGACAGACGCTCTGATGCCGGTGCCATGGCCTGAAAGCTCCCTAA	1305			
QY	268	AAACCTCTGGAGGAGGCGGTGTGAGGGCGTCTGACCAAGTGAATCCAGCCCTCG	327	1251	TAGACTTTGCTGGTACCCTGCAAGGAGCATTTCCACGGAGTCTGTACCATCATTTCTG	1310			
DB	286	AAACCTCTGGAGGAGGCGGTGTGAGGGCGTCTGACCAAGTGAATCCAGCCCTCG	345	1306	TAGACTTTGCTGGTACCCTGCAAGGAGCATTTCCACGGAGTCTGTACCATCATTTCTG	1365			
QY	328	TGGGCTCAAAATATAACATTTGGGTGAACCTGATTTCCCTAGATGCCAAAAGGAGATG	387	1311	ACCCACCTGGAGATCACCCAGAAACACAGTCTGCGAGCCCTGTGGATGTGGATGAGATGT	1370			
DB	346	TGGGCTCAAAATATAACATTTGGGTGAACCTGATTTCCCTAGATGCCAAAAGGAGATG	405	1366	ACCCACCTGGAGATCACCCAGAAACACAGTCTGCGAGCCCTGTGGATGTGGATGAGATGT	1425			
QY	388	CCAATGGCAACATAGTCTATGAGAACTGAGAAATGAGGCTGGTGTATCTGCTGATC	447	1371	GTCTGCTGACTGTGAGACGAACTTCAATGGGTCTGGGACGTAATCTGTGAACTCACCC	1430			
DB	406	CCAATGGCAACATAGTCTATGAGAACTGAGAAATGAGGCTGGTGTATCTGCTGATC	465	1426	GTCTGCTGACTGTGAGACGAACTTCAATGGGTCTGGGACGTAATCTGTGAACTCACCC	1485			
QY	448	CATATGTTTAACTGGACAGCATGGTCCAGAGGACAGTGAACGGGAAATGGCACCGCC	507	1431	TGGGGGATGACAAAGCCTGGCTCTCAGAGACCCCTGATTTCTGTCTCTGACAGAGACC	1490			
DB	466	CGTATGTTTAACTGGACAGCATGGTCCAGAGGACAGTGAACGGGAAATGGCACCGCC	525	1486	TGGGGGATGACAAAGCCTGGCTCTCAGAGACCCCTGATTTCTGTCTCTGACAGAGACC	1545			
QY	508	AAAGCCATCAACAGTCTCCCTGATGGGAAACCTTTTCTCACCACCCCGGATGGAGAA	567	1491	CAGCCTCGCCTTTTAAAGGATGGCAAAACAGTGCCTGATCTCCGTGGCTGTGGCCATAT	1550			
DB	526	AAAGCCATCAACAGTCTCCCTGATGGGAAACCTTTTCTCACCACCCCGGATGGAGAA	585	1546	CAGCCTCGCCTTTTAAAGGATGGCAAAACAGGCCCCCTGATCTCGITGGCTGTGGCCATAT	1605			
QY	568	GATGGAATTTCACTACGCTCTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGAT	627	1551	TTGTCACTGTGATCTCCCTTTGGTGTACAAAAACAAAGGAATACAACCCCAATAGAAA	1610			
DB	586	GATGGAATTTCACTACGCTCTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGAT	645	1606	TTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACAAAGGAATACAACCCCAATAGAAA	1665			
QY	628	GTTCACTGAGAGTTTCTGTGAACACAGCCAAATGTGACACTTTGGGCTCAACTCATGGAAG	687	1611	ATAGTCTGGGAAATGTGGTTCAGAAAGCAAGGCTCGAGTGTCTTCTCAACCGTGCAAAAG	1670			
DB	646	GTTCACTGAGAGTTTCTGTGAACACAGCCAAATGTGACACTTTGGGCTCAACTCATGGAAG	705	1666	ATAGTCTGGGAATGTGGTTCAGAAAGCAAGGCTCGAGTGTCTTCTCAACCGTGCAAAAG	1725			
QY	688	TGACTGTCTACAGAAACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAAGATGTGT	747	1671	CGGTGTTCTCCCGGAAACACAGGAAAGGATCGGCTACTCAAAAAACCAAGAAATTTAAAG	1730			
DB	706	TGACTGTCTACAGAAACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAAGATGTGT	765	1726	CGGTGTTCTCTCCCGGAAACACAGGAAAGGATCGGCTACTCAAAAAACCAAGAAATTTAAAG	1785			
QY	748	ACGTGGTAAACAGATTCAGTCTGTTGTGACTATGTTCCAGAAAGACGATCGAAAT	807	1731	GAGTTTCTTAAATTTGACACCTTGTGTTCTGAAAGCTCACTTTTTCAGTGCCATTTGATGTGAGA	1790			
DB	766	ACGTGGTAAACAGATTCAGTCTGTTGTGACTATGTTCCAGAAAGACGATCGAAAT	825	1786	GAGTTTCTTAAATTTGACACCTTGTGTTCTGAAAGCTCACTTTTTCAGTGCCATTTGATGTGAGA	1845			
QY	808	CATCCGACGAAACCTTCC-CAAAGATCTCCCAATATGTTTGTGATGCTGATTCATGATC	866	1791	TGTGCTGGAGTGGCTTATTAAACCTTTTTTCCCTAAAGATTAATGTTTAAATAGATATTGTGG	1850			
DB	826	CATCCGACGAAACCTTCC-CAAAGATCTCCCAATATGTTTGTGATGCTGATTCATGATC	885	1846	TGTGCTGGAGTGGCTTATTAAACCTTTTTTCTCTAAAGATTAATGTTTAAATAGATATTGTGG	1905			
QY	867	CTAGCCACTTCTCAATTTATCTACCAATTAACCTACAGTGGAGCTTGGGGATAATACTG	926	1851	TTTGGGGAAGTTGAAATTTTTTATAGTTTAAATGTCAATTTTAGAGATGGGAGAGGATTA	1910			
DB	886	CTAGCCACTTCTCAATTTATCTACCAATTAACCTACAGTGGAGCTTGGGGATAATACTG	945	1906	TTTGGGGAAGTTGAAATTTTTTATAGTTTAAATGTCAATTTTAGAGATGGGAGAGGATTA	1965			
QY	927	GCTGTTTGTGTTCCACCAATCACTATGTAATCACAGTATGCTCAATGGAACCTTCA	986	1911	TACTGAGGACGCTTCAGCCATGTTGTGAAACCTGATATAAAGCAACTTAGCAAGGCTTCTT	1970			
DB	946	GCCTGTTTGTGTTCCACCAATCACTATGTAATCACAGTATGCTCAATGGAACCTTCA	1005	1966	TACTGAGGACGCTTCAGCCATGTTGTGAAACCTGATATAAAGCAACTTAGCAAGGCTTCTT	2025			
QY	987	GCCTTAACTCACTGTGAAAGCTGACGACCAAGGACCTTGTCCGCGACCGCCACCAACCAC	1046	1971	TTCAATTTTTTATGTTTCACTTATAAGTCTTAGGTAACTAGTGTAGATAGAAACAATG	2030			
DB	1006	GCCTTAACTCACTGTGAAAGCTGACGACCAAGGACCTTGTCCGCGACCGCCACCAACCAC	1065	2026	TTCAATTTTTTATGTTTCACTTATAAGTCTTAGGTAACTAGTGTAGATAGAAACAATG	2085			
				2031	TGTCCCGAGAGTAGGAGAGAGAGCTACTATTGATTAGAGCCTTAACCCAGGTTAACTGCAA	2090			
				2086	TGTCCCGAGAGTAGGAGAGAGAGCTACTATTGATTAGAGCCTTAACCCAGGTTAACTGCAA	2145			
				2091	GAAGAGGCGGGATACCTTTTCAGCTTTTCCATGTTAACTGTATGTGATATAAGCCAATGTAGTCCA	2150			

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: September 10, 2004, 10:01:28 ; Search time 40.5 Seconds  
(without alignments)  
6842.404 Million cell updates/sec

Title: US-10-039-272-1

Perfect score: 4850

Sequence: 1 cggcagcaggccagagga.....aaaaaaaaaaaaaaaaaaaaa 2661

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/cgn2 1/USPTO spool p/US10039272/runat\_07092004\_144222\_21619/app\_query.fasta\_1.2823  
-DB=SwissProt 42 -QFMT=fastan -SUFFIX=rs -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USRP=US10039272 @CGN 1.1 48 @runat\_07092004\_144222\_21619 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEOUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=1 -XGAPOP=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt 42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2999	61.8	560	1 NMB_HUMAN	Q14956 homo sapien
2	1563.5	32.2	559	1 QNR_COTJA	Q90372 ccturnix co
3	635	13.1	626	1 PM17_MOUSE	Q60696 mus musculus
4	617	12.7	762	1 P115_CHICK	Q98917 gallus gall
5	593	12.2	661	1 PM17_HUMAN	P40967 homo sapien
6	436.5	9.0	491	1 PM17_BOVIN	Q06154 bos taurus
7	129	2.7	926	1 PTN4_HUMAN	P29074 homo sapien
8	116.5	2.4	555	1 ODP2_RAT	P08461 rattus norv
9	115.5	2.4	1093	1 SM5B_MOUSE	Q60519 mus musculus
10	111	2.3	525	1 HRG_HUMAN	P04196 homo sapien
11	111	2.3	5038	1 PCLO_MOUSE	Q9qvx7 mus musculus
12	110.5	2.3	1874	1 POLR_KYUVJ	P36304 kenedya ye
13	110	2.3	1225	1 CTD2_HUMAN	Q9ubc3 homo sapien
14	109.5	2.3	5703	1 MUB3_HUMAN	Q9h84 homo sapien
15	106.5	2.2	427	1 IRE3_HUMAN	Q14653 homo sapien
16	106	2.2	1537	1 FLOI_HUMAN	P32768 saccharomyc
17	104	2.1	435	1 GAT3_XENLA	P23773 xenopus lae
18	104	2.1	4303	1 PKD1_HUMAN	P98161 homo sapien

19	103.5	2.1	5147	1 PCLO_HUMAN	Q9y6v0 homo sapien
20	103	2.1	626	1 GPBA_HUMAN	P07359 homo sapien
21	103	2.1	921	1 RB_MOUSE	P13405 mus musculus
22	103	2.1	1009	1 YE68_METJA	Q58863 methanococc
23	102.5	2.1	579	1 SOC6_MOUSE	Q8vht2 mus musculus
24	102.5	2.1	979	1 RFX1_HUMAN	P22670 homo sapien
25	102	2.1	1213	1 FMN_CHICK	Q05858 gallus gall
26	101	2.1	2090	1 N214_HUMAN	P35658 homo sapien
27	100	2.1	1068	1 DAM2_HUMAN	Q86565 homo sapien
28	99.5	2.1	1587	1 LMG3_HUMAN	Q9y6n6 homo sapien
29	99	2.0	450	1 AT10_MOUSE	P58459 mus musculus
30	99	2.1	499	1 CH1B_SERMA	P11797 serratia ma
31	98.5	2.0	594	1 RYK_MOUSE	Q01887 mus musculus
32	98	2.0	1059	1 CAPU_DROME	Q24120 drosophila
33	98	2.0	4548	1 APOA_HUMAN	P08519 homo sapien
34	98	2.0	5120	1 PCLO_CHICK	Q9pu36 gallus gall
35	97.5	2.0	785	1 CAD7_HUMAN	Q9ulb5 homo sapien
36	97.5	2.0	1858	1 P3K2_DICDI	P54674 dictyosteli
37	97	2.0	340	1 CYSP_SCHMA	P25792 schistosoma
38	97	2.0	673	1 REP_ECOLI	P09980 escherichia
39	96.5	2.0	488	1 BTB1_MOUSE	P58544 mus musculus
40	96	2.0	377	1 AR41_SCHPO	P78774 schizosacch
41	96	2.0	512	1 PER1_VOLCA	P81131 volvox cart
42	95.5	2.0	671	1 ENV_FENV1	P31791 feline endo
43	95.5	2.0	1161	1 DAN4_YEAST	P47179 saccharomyc
44	95.5	2.0	1247	1 SAH1_HUMAN	Q94685 homo sapien
45	95.5	2.0	1620	1 HED1_HUMAN	Q9ult8 homo sapien

## ALIGNMENTS

RESULT 1  
NMB\_HUMAN  
ID AC Q14956; STANDARD; PRT; 560 AA.  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, last sequence update)  
DT 15-MAR-2004 (Rel. 43, last annotation update)  
DE Putative transmembrane protein NMB precursor (Transmembrane glycoprotein HGFIN)  
GN GNMB OR NMB OR HGFIN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RN SEQUENCE FROM N.A.  
RC TISSUE=Melanoma;  
RX MEDLINE=95113576; PubMed=7814155;  
RA Waterman M.A.J., Ajubi N., van Dinter I.M.R., Degen W.G.J., van Muijen G.N.P., Ruiter D.J., Bloemers H.P.J.;  
RT "NMB, a novel gene, is expressed in low-metastatic human melanoma cell lines and xenografts.";  
RL Int. J. Cancer 60:73-81(1995).  
[2]  
RN SEQUENCE FROM N.A.  
RC TISSUE=Peripheral blood;  
RX MEDLINE=22498106; PubMed=12609765;  
RA Bandari P.S., Qian J., Yehia G., Joshi D.D., Maloof P.B., Potian J., Oh H.S., Gascon P., Harrison J.S., Rameswar P.;  
RT "Hematopoietic growth factor inducible neurokinin-1 type: a transmembrane protein that is similar to neurokinin 1 interacts with substance P.";  
RL Regul. Pept. 111:169-178(2003).  
CC -!- FUNCTION: Could be a melanogenic enzyme (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
CC -!- TISSUE SPECIFICITY: Not restricted to the melanocytic lineage.  
CC -!- DEVELOPMENTAL STAGE: Expression in poorly metastatic melanoma cell lines; no expression in highly metastatic melanoma cell lines.  
CC -!- SIMILARITY: BELONGS TO THE PMEL-17/NMB FAMILY.  
CC -!- SIMILARITY: Contains 1 PKD domain.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; X76534; CAA54044.1; -;  
DR EMBL; AF322909; AAG42839.1; -;  
DR PIR; I38065; I38065.  
DR Genew; HGNC:4462; GENMB.  
DR MIM; 604368;  
DR GO; GO:0016021; C: integral to membrane; TAS.  
DR GO; GO:0008285; P: negative regulation of cell proliferation; TAS.  
DR InterPro; IPR000501; PKD.  
DR Pfam; PF00801; PKD; 1.  
DR SMART; SM00089; PKD; 1.  
DR PROSITE; PS50093; PKD; 1.  
KW Signal; Transmembrane; Glycoprotein; Polymorphism.  
FT SIGNAL 1 21 POTENTIAL.  
FT CHAIN 22 560 PUTATIVE TRANSMEMBRANE PROTEIN NMB.  
FT DOMAIN 22 486 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 487 507 POTENTIAL.  
FT DOMAIN 508 560 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 240 327 PKD.  
FT DOMAIN 320 332 POLY-PRO.  
FT CARBOHYD 93 93 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 134 134 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 146 146 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 200 200 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 249 249 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 275 275 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 296 296 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 300 300 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 306 306 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 312 312 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 447 447 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 455 455 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT VARIANT 195 195 S -> C (in dbSNP:530436).  
FT VARIANT 197 197 /FTID=VAR 012076.  
FT VARIANT 197 197 N -> H (in dbSNP:530413).  
FT /FTID=VAR 012077.  
SQ SEQUENCE 560 AA; 62643 MW; 570035B48CE3BCC CRC64;

Alignment Scores:  
pred. No.: 3,37e-248 Length: 560  
Score: 2999.00 Matches: 560  
Percent Similarity: 99.82% Conservative: 0  
Best Local Similarity: 99.82% Mismatches: 1  
Query Match: 61.84% Indels: 1  
DB: 1 Gaps: 0

US-10-039-272-1 (1-2661) x NMB\_HUMAN (1-560)

QY 60 ATGGAATGCTCTACTATTCTCTGGGATTTCTGCTGCTGCTGCAAGATTGCCACTTGAT 119

Db 1 MetGluCysLeuTyrPheLeuGlyPheLeuLeuLeuAlaAaArgLeuProLeuAsp 20

QY 120 GCCGCCAAACGATTTCATGATGCTGCGCAATGAAAGACCTTCTGCTTACATGAGGAG 179

Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40

QY 180 CACAATCAATTAATGCTGCTCTGATGAAATGCTGAAATGAAATGAAATGAAATGAAATG 239

Db 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60

QY 240 GTCTGGAGCGGGAGACATGAGTGGGAAACCTCTGGAAGGAGGCGCGTGTGACGCG 299

Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla 80

QY 300 GTCCTGACGAGTCACTACCGCCCTCTGGGGCTCAATATAAATTTGGCGGTGAACCTG 359

Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100

QY 360 ATATTCCTAGATGCCAAAAGAGATGCCAATGGGCAACATAGTCTATGAGAGAACTGC 419  
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120  
QY 420 AGAAATGAGCTGGTTTATCTGCTGATCCATATGTTTACAACTGGACAGCATGTGTCAG 479  
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTyrSerGlu 140  
QY 480 GACAGTCACGGGAAAATGGCACCGGCAAGCCATCATACGCTTCTCCCTGATGGGAAA 539  
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160  
QY 540 CCTTTTCTCACCACCCCGATGGAGATGGAATTTTCATCTACGCTTCCACACACTT 599  
Db 161 ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180  
QY 600 GGTCAATGATTTCCAGAAATTTGGGACGATGTTTCAGTGAGAGTTTCTGTGCAACACAGCCCAAT 659  
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200  
QY 660 GTGACACTTGGGCTCAACTCATGGAAGTGTCTTACAGAAGACATGAGCGGCAAT 719  
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220  
QY 720 GTTCCCATCGCACAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCCTGTGTGTGTG 779  
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240  
QY 780 ACTATGTTCCAGAGAAAGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838  
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260  
QY 839 ATTATGTTGTGCTCCTGATTCATGATCCTACCCACTTCTCCTCAATATTCTACCATTAAC 898  
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280  
QY 899 TACAAGTGGAGCTTCGGGATTAATCTGCGCTGTTGTTGTTCCACCAATCATCTGTGAAT 958  
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300  
QY 959 CACAGTATGCTCAATGGAACTTCACTGCTTAACCTCACTGTGAAAGCTGCAGACCA 1018  
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320  
QY 1019 GGACCTTGTCCGCCACCGCCACACACACACACCTTCAAAACCCACCTCTTTAGGA 1078  
Db 321 GlyProCysProProProProProProProProProProProProProProProSerLeuGly 340  
QY 1079 CTTGCTGGTGACAAACCCCTGAGCTGAGTAGGATTCCTGATGAAATGCTCCAGATTAAC 1138  
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnGlnIleAsn 360  
QY 1139 AGATATGGCCACTTCAAGCCACCATCACAATGTAGAGGAATCTTAGAGTTAAATC 1198  
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyLeuLeuValAsnIle 380  
QY 1199 ATCCAGATGACACACGCTCTGATGCGGTCGCCCTGAGGCTGAAAGCTCCCTTAATAGACTTT 1258  
Db 381 IleGlnMetThrAspValLeuMetProValProProProProProProProProProPro 400  
QY 1259 GTCGTGACCTGCCAAGGAGCATTCACCGAGGTCTGTACCATCATTTCTGACCCACCC 1318  
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420  
QY 1319 TGCAGATATCCACAGAACACAGTCTGCAGCCCTGTGGATGTGGATGTGGATGTGGATGTGCTGT 1378  
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440  
QY 1379 ACTGTGAGAGAACTTCAATGGGTCTGGGACGTACTGTGTGGAACCTCACCCTGGGGAT 1438  
Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460



Db 227 ThrThrTyrValValThrAspGlyLeuProIleLeuValSerMetSerGlnLysHisAsp 246  
QY 801 CGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCCATTATGTTTGTATGCTCTGATT 859  
Db 247 ArgAsnIleSerAspSerIlePheIleLysAspSerProIleThrPheAspValLysIle 266  
QY 860 CATGATCCTAGCCATCTCTCAATATTTACCAATTAACCTACAACTGAGCTCGGGAT 919  
Db 267 HisAspProSerTyrTyrLeuAsnAspSerAlaIleSerTyrLysTrpAsnPheGlyAsp 286  
QY 920 AATACTGCGCTGTTGTTTCCACCAATCATCTGATACACAGTATGCTGCTCAATGGA 979  
Db 287 GlySerGlyLeuPheValGluSerGlyAlaThrThrSerHisThrPheSerLeuGlnGly 306  
QY 980 ACCTTCAGCCTTAACCTCACTGTGAAAGCTGACAGCAGACCTGTGTCGCCACCGCCA 1039  
Db 307 AsnPheThrLeuAsnLeuThrValGlnAlaIleProValProCys-----LysPro 324  
QY 1040 CCACACCCAGCCTTCAAAACCCACCCCTCTTTA----- 1075  
Db 325 ValThrProThrProSerLeuProThrProAlaValThrThrAspAlaSerSerAsnSer 344  
QY 1076 GGACCTGCTGTGCAACCCCTCGAGCTGAGTAGATTCTGTGATGAAACTGCCAGATT 1135  
Db 345 AspProSerAlaProAsnGluMetAlaGluAspAsn---ProAspGlyGlyCysHisIle 363  
QY 1136 AACAGATATGCGCACTTTCAAGCCACCATCAATAGTAGAGGGAATCTTAGAGTTAAAC 1195  
Db 364 TyrArgTyrGlyTyrThrAlaGlyIleThrIleValGluGlyIleLeuGluValAsn 383  
QY 1196 ATCATCCAGATGACAGAGCTCTGATGCGGCTGCGCATGGCTGAAAGCTCCCTATAGAC 1255  
Db 384 IleIleGlnMetThrSerIleGlnMetThrGluSerGlnAlaGluAsnProLeuValAsp 403  
QY 1256 TTTGTGCTGACCTGCAAGGAGCATTCCTCCAGGAGTCTGTACCATCATTTTCTGACCCC 1315  
Db 404 PheValValThrCysGlnGlySerPheProThrAspValCysThrAlaValSerAspPro 423  
QY 1316 ACCTCGAGATCACCAGAACACAGTCTGACGCCCTGTGGATGTGATGAGATGTGTCTG 1375  
Db 424 ThrCysGlnValSerGlnGlyMetValCysAspProValValValThrAspGluCysVal 443  
QY 1376 CTGACTGTGACGACCACTTCAATGGTCTGGAGCTGACTGTGAACTCACCCTGGGG 1435  
Db 444 LeuThrIleArgArgAlaPheAspGluProGlyThrTyrCysIleAsnIleThrLeuGly 463  
QY 1436 GATGACACAAAGCTGCTCTCACGAGCACCTGATTTCTGTTCTCTGACAGACCCAGCC 1495  
Db 464 AspAspThrSerGlnAlaLeuAlaSerAlaLeuIleSerVal-----AsnGlyGly 480  
QY 1496 TCGCCTTTAAGATGCAACAGTCCCTGATCTCGTTGGCTGTGCTGCTGCTGCTGCTG 1555  
Db 481 SerSerSerGlyThrThrIysGlyValPheIlePheLeuGlyLeuLeuAlaValPheGly 500  
QY 1556 ACTGTGATCTCCCTCTGTGTGTACAAAACACAGGAATACAAACCAATAGTAATAGT 1615  
Db 501 AlaIleGlyAlaPheValLeuTyrIysArgTyrLysGlnIleThrLysProIleGluArgSer 520  
QY 1616 CTTGGGAATGTGTGCAAGCAAGGCTGAGTGTCTTTCTCAACCGTGCAGAAAGCCGTG 1675  
Db 521 AlaGlyGlnAlaGluAsnGlnGluGlyLeuSerAlaTyrValSerAsnPheIysAlaPhe 540  
QY 1676 TTCTTCCCGGAAACAGGAAAGATTCGCTACTCAAAACCAA 1720  
Db 541 PhePheProLysSerThrGluArgAsnProLeuLeuLysSerLys 555

RESULT 3

PM17 MOUSE STANDARD; PRT; 626 AA.  
ID PM17 MOUSE  
AC Q60696;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Melanocyte protein Pmel 17 precursor (Silver locus protein).  
GN STLV OR PMEL17 OR D10H12S53E OR SI.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
SEQUENCE FROM N.A.  
RF STRAIN=C57BL/6; TISSUE=Skin;  
RC MEDLINE=95175358; PubMed=7870580;  
RA Kwon B.S., Halaban R., Ponnazhagan S., Kim K., Chintamaneni C.,  
RA Bennett D., Pickard R.T.;  
RT "Mouse silver mutation is caused by a single base insertion in the  
RT putative cytoplasmic domain of Pmel 17.";  
RL Nucleic Acids Res. 23:154-158(1995).  
CC -!- FUNCTION: Could be a melanogenic enzyme.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
CC -!- TISSUE SPECIFICITY: Preferentially expressed in melanocytes.  
CC -!- DISEASE: DEFECTS IN SILV ARE THE CAUSE OF THE SILVER COAT COLOR  
CC WHICH SEEMS TO BE DUE TO PREMATURE DEATH OF PIGMENT CELLS DURING  
CC THE HAIR CYCLE.  
CC -!- SIMILARITY: BELONGS TO THE PMEL-17/NMB FAMILY.  
CC -!- SIMILARITY: Contains 1 PKD domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: U14133; AAA69538.1; --  
CC PIR: S53871; S53871.  
CC MGD: MGI:98301; Si.  
DR InterPro: IPR000601; PKD.  
DR Pfam: PF00801; PKD; 1.  
DR SMART: SM00089; PKD; 1.  
DR PROSITE: PS50093; PKD; 1.  
KW Transmembrane; Glycoprotein; Signal; Melanin biosynthesis; Repeat;  
KW Disease mutation.  
KW SIGNAL 1 24  
FT CHAIN 25 626  
FT DOMAIN 25 562  
FT TRANSMEM 563 583  
FT DOMAIN 584 626  
FT DOMAIN 255 292  
FT DOMAIN 315 327  
FT REPEAT 328 340  
FT REPEAT 341 353  
FT REPEAT 354 366  
FT REPEAT 367 379  
FT REPEAT 380 392  
FT REPEAT 393 411  
FT CARBOHYD 81 81  
FT CARBOHYD 106 106  
FT CARBOHYD 111 111  
FT CARBOHYD 535 535  
FT VARIANT 170 170  
FT VARIANT 175 175  
FT VARIANT 373 373  
FT VARIANT 471 471  
FT VARIANT 603 626  
SQ SEQUENCE 626 AA; 65980 MW; 7AB941D2E3FB1044 CRC64;  
Alignment Scores:  
Pred. No.: 4,88e-46 Length: 626  
Score: 635.00 Matches: 184  
Percent Similarity: 41.45% Conservative: 97  
Best Local Similarity: 27.14% Mismatches: 207

Query Match:	13.09%		Indels:	191
DB:	1		Gaps:	22
 US-10-039-272-1 (1-2661) x PM17_MOUSE (1-626)				
QY	78	TTCTCGGATTCCTGCTCTGGCTGCAAGATTCCACTTGTATGCC-----GCCAAACGA	131	
Db	9	PheLeuProValLeuValleuSerAlaLeuLeuAlavalGlyAlaLeuGluGlySerArg	28	
QY	132	TTTCATGATGCTGCTGGCAATGAAGAACCCTTGCTGTACATGAGGAGCACAAATCAATTA	191	
Db	29	AsnGlnAspTrpLeuGlyValProArgGlnLeuVal-----	40	
QY	192	AATGGCTGCTCTCTGATCAAAAATGACTTGAATGAAAAAATCTTACCAGTGTGGAACGG	251	
Db	41	-----ThrIysThrTpAsnArgGlnLeuTrpThr-----	53	
QY	252	GGAGCATGAGTGGAAAAACTCCTGGAAGGAGGCCTGTGAGCGCTCTGACCACT	311	
Db	54	--GluValGlnGlySerAsnCysTrpArgGlyGlyGlnValSerLeuArgVallleAsn	72	
QY	312	GACTCACACAGCCCTGCTGGGCTCAAATAACAATTGCGGTGAACATGATATCCCTAGA	371	
Db	73	AspGlyProThrLeuValGlyAlaAsnLaserPheSerIleAlaLeuHisPheProGly	92	
QY	372	TGCCAAAAGGAAGATGCCAATGCCACATAGTAGTCTATGAGAAAGAC---TGCAGAATGAG	428	
Db	93	SerGlnLysValLeuProAspspGlyGlnValIleTrpAlaAsnAsnThrIleileAsnGly	112	
QY	429	GCTGTTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGTTCAGAGACAGTGAC	488	
Db	113	SerGlnValTrpGlyGlyGlnProValtyrPro-----GlnGluProAspAsp	128	
QY	489	GGGNAANAATGGCACCGGCCAAAGCCATCANACGTTCTCCCTGATGGGAACCTTTTCT	548	
Db	129	Ala-----CysValPheProAspGlyGlyProCysPro	139	
QY	549	CACCACCCCGGATGAGAGAATGGAATTCATCTACGTCTTCCACACACTTGGTCAGTAT	608	
Db	140	SerGlyProLysProProLyArgSerPheValTyrrValTrpLysThrTrpGlyLysTyr	159	
QY	609	TTCCAGAAATTGGGACGATGTTCTAGTGAGAGTTTCTGTGAACACAGCCAATGTGACATT	668	
Db	160	TrpGlnValLeuGlyGlyProValSerArgSerSerIleAlaThrArgHISalalysLeu	179	
QY	569	GGGCTCACTCATGGAAGTGACTGCTCPACAGAAGACATGGA---CGGGCATATGTTCCC	725	
Db	180	GlyThrHisThrMetGluValThrValtyrHisArgArgGlySerGlnSerTyrValPro	199	
QY	726	ATCCCAACAAGTGAAGATGTCAGTGGTACACATCAGATTCCTGTGTTTGTGCACTATG	785	
Db	200	LeuAlaHisAlaSerSerThrPheThrPheThrIleThrAspGlnValProPheSerValSer	219	
QY	786	TTCCAGAA-----GAACGATCGAAATTCATCCCGACGAACCTTCCCCAA	829	
Db	220	SerGlnLeuGlnAlaLeuAspGlyGluThrLysHisPheLeuArgAsnHis-----	236	
QY	830	GATCTCCCATATGTTTGATGTCCTGATTCATGATCTAGCCACTTCTCCTCAATTATCT	889	
Db	237	-----ProLeuIlePheAlaLeuGlnLeuHisAspProSerGlyTyrLeuAlaGluAla	254	
QY	890	ACCATTAACTACAAGTGAGCTTCGGGGATAATACTGGCCCTGTTGTTTCCACCAATCAT	949	
Db	255	AspLeuSerTyrThrTrpAspPheGlyAspGlyThrGlyThrLeuIleSerArgAlaLeu	274	
QY	950	ACTGTGAATCACACGTATGCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAGCT	1009	
Db	275	AspValThrHisThrTyrLeuGluSerGlySerValThrAlaGlnValValLeuGlnAla	294	
QY	1010	GCA-----	1012	
Db	295	AlaIleProLeuValSerCysGlySerSerProValProGlyThrThrAspGlyTyrMet	314	



DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE 115 kDa melanosomal matrix protein precursor.  
GN WNP115.

OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.

OC NCBI\_TaxID=9031;  
RN [1]  
RN SEQUENCE FROM N.A.

RP STRAIN=White leghorn; TISSUE=Retinal pigment epithelium;  
RC MEDLINE=98311098; PubMed=3409326;  
RX MEDLINE=92020667; PubMed=1924173;

RA Mochii M., Agata K., Eguchi G.;  
RA "Complete sequence and expression of a cDNA encoding a chicken  
RT 115-kDa melanosomal matrix protein.";

RL Pigment Cell Res. 4:41-47(1991).  
RN [2]

RP CHARACTERIZATION  
RC STRAIN=White leghorn; TISSUE=Retinal pigment epithelium;  
RX MEDLINE=98311098; PubMed=3409326;

RA Mochii M., Agata K., Kobayashi H., Yamamoto T.S., Eguchi G.;  
RA "Expression of gene coding for a melanosomal matrix protein  
RT transcriptionally regulated in the transdifferentiation of chick  
RT embryo pigmented epithelial cells.";

RL Cell Differ. 24:67-74(1988).  
CC -!- FUNCTION: Might be required for polymerization of melanin onto the  
CC core structure of melanosomes with enzymic function of tyrosinase.

CC -!- SUBCELLULAR LOCATION: ON THE FIBROUS MATRIX STRUCTURE OF THE  
CC PREMELANOSOME.

CC -!- TISSUE SPECIFICITY: Specific to pigmented epithelial cells and  
CC melanocytes. Not expressed in lens, neural retina, brain, heart,  
CC gizzard or liver.

CC -!- DEVELOPMENTAL STAGE: Expressed during the redifferentiation of  
CC pigmented epithelial cells (PEC).

CC -!- PTM: Glycosylated.

CC -!- SIMILARITY: BELONGS TO THE PMEL-17/NMB FAMILY.

CC -!- SIMILARITY: Contains 1 PKD domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; D88348; BAAL3589.1; -  
DR InterPro; IPR000601; PKD.  
DR Pfam; PF00801; PKD; 1.

DR SMART; SM00089; PKD; 1.  
DR PROSITE; PS00993; PKD; 1.

KW Signal; Glycoprotein; Repeat.  
FT SIGNAL 1 19  
FT CHAIN 20 762  
FT DOMAIN 223 323

FT REPEAT 441 532  
FT REPEAT 441 464  
FT REPEAT 465 488

FT REPEAT 489 508  
FT REPEAT 509 532  
FT CARBOHYD 111 111

FT CARBOHYD 115 115  
FT CARBOHYD 346 346  
FT CARBOHYD 651 651

FT CARBOHYD 659 659  
SQ SEQUENCE 762 AA; 77356 MW; 172C8DDB4FCE7C6 CRC64;

POTENTIAL.  
115 KDA MELANOSOMAL MATRIX PROTEIN.  
PKD.  
4 X 20-24 AA APPROXIMATE TANDEM REPEATS.

N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).

Alignment Scores:  
Pred. No.: 1,81e-44 Length: 762  
Score: 617.00 Matches: 171  
Percent Similarity: 35.60% Conservative: 80  
Best Local Similarity: 24.26% Mismatches: 186

Query Match: 12.72% Indels: 269  
DB: 1 Gaps: 15  
US-10-039-272-1 (1-2661) x P115\_CHICK (1-762)  
QY 90 CTCTCTCTGCTCAAGATTGCCACATTCATGTCGCGCAACAGATTTCATGATGTGCTGGC 149  
DB 8 ValLeuLeuAlaLeuLeuAlaLeu-----ValThrAla 19  
QY 150 AATGAAAGACCTTCTGCTTACATGAGGAGCACAATCAATTAATGCTGTCT----- 203  
DB 20 GlnGlnArgGlyGlyArgSerArgGlyValLysGlySerAlaTyrGlyArg 39  
QY 204 ---TCTGATGAAATGATCGAATGAAACTCTACCCAGTGTGCAAGCGGGAGACATG 260  
DB 40 ProAlaProPheArgSerTyrPheThrAlaArgTyrArgProThrGlnGluGlyThrAla 59  
QY 261 AGTGGGAAACCTCTGGAAGGAGCGGTGTCAGCGGTCTCTGACGACGACTCAACCA 320  
DB 60 ArgGlnAsnAspCysTyrPheGlyGlyAspValThrPheAspIleSerAsnAspAlaPro 79  
QY 321 GCCCTCTGCTGCTCAAAATATAACATTTGCGGTGAACCTGATATTCCTAGATGCAAAAG 380  
DB 80 ThrLeuValGlyAlaArgAlaThrPheSerIleAlaLeuArgPheProGlyThrGlnThr 99  
QY 381 GAAGATGCCAATGCAACATAGTCTATGAGAAAGACTGC-----AGAAATGAGCTGGT 434  
DB 100 ValLeuProAspGlyArgValValTyrSerGlnAsnCysThrValAsnGlyThrArgMet 119  
QY 435 TTATCTGCTGATTCATATGTTTCAACTGAGCAGCATGTCAGAGGACAGTGCAGCGGAA 494  
DB 120 LeuGlnGlyAspPro---ValTyrPro---GluGlnLeuAlaGluGlySerAspGly--- 136  
QY 495 AATGACACCGGCCAAAGCCATCAATGCTTCTCCCTGATGGGAAACCTTTCTCCACAC 554  
DB 137 -----ValPheProAspGlyGlnProPheProArgSer 147  
QY 555 CCCGATGAGAGAGATGGAATTTCTATCTACGCTCTCCACACATTTGTCAGTATTTCCAG 614  
DB 148 AlaTyrGlyLysArgGlyArgPheValTyrValTyrThrTyrGlyArgTyrThrGln 167  
QY 615 AAATTTGGACGATGTTCACTGAGAGTTCCTGTAACACAGCAATGACATCTGGGCT 674  
DB 168 ValValAspGlyAlaThrSerGlnLeuThrValGlyThrAspGlyValAlaLeuGlySer 187  
QY 675 CAACCTCATGGAAGTCACTGTCTACAGAAGACATGACGCG---GCATATGTTCCCATCGCA 731  
DB 188 TyrThrMetGluValValValTyrHisTyrArgGlyArgGlnArgPheIleProIleGly 207  
QY 732 CAAGTGAAGATGTGTAGTGTGTAACATCAGATTCCTGCTGTTGTGACTATGTTCCAG 791  
DB 208 HisAlaSerThrGlnPheSerIleThrAspGlnValProIleAlaValAspValThrGln 227  
QY 792 AA-----GAACGATCGAAATTCATCCGACGAAACCTTCCCAAGATCTC 835  
DB 228 LeuGluValAlaAlaGlyAspGlyGlySerPheValArgAsnArg----- 242  
QY 836 CCCATTATGTTGATGTCCTGATTCATGATCTAGCCATCTCTCAATTTCTTACCATT 895  
DB 243 ProValAlaPheAsnValArgLeuHisAspProSerHisTyrLeuArgAspAlaAspIle 262  
QY 896 AACTCAAGTGCAGCTTCGGGGATATATCTGCGCTGTTGTTTCCCAATCATCTG 955  
DB 263 SerTyrSerTyrAspPheGlyAspGlnSerGlyThrLeuIleSerArgSerProThrVal 282  
QY 956 AATCAGCATGTGTGCTCAATGAACTTCAGCTTTAACTCACTGCTGAAAGCTGACGCA 1015  
DB 283 ThrHisThrTyrLeuGlnAlaGlySerPheAlaAlaArgLeuValLeuGlnAlaAla 302  
QY 1016 CCA-----GGACCTTGTCCGCCACCCACCCACCCACCCACCTTCAAAACCC 1063  
DB 303 ProLeuSerSerCysGlyThrSerAlaProProValValAspPro---ThrThrGlyPro 321

```

QY 1064 ACCCTCTCTTAGGACCTGCTGTCGACACCC----- 1096
Db 322 ValProSerLeuGlyProThrAlaThrGlnProValGlyProThrGlySerGlyThrAla 341
QY 1096 ----- 1096
Db 342 ThrAlaProSerAsnLeuThrGlySerGlyThrAlaAlaAlaProGlyThrThrAlaAla 361
QY 1096 ----- 1096
Db 362 ProArgAlaSerGlyAlaProAlaGluProThrGlyValSerValaValaValuSerAsp 381
QY 1097 ---CTGAGCTGAGTAGGATTCCTGAT----- 1120
Db 382 SerMetGluThrGluProLeuProAspProValLeuSerThrAlaValaAlaAspAlaAla 401
QY 1120 ----- 1120
Db 402 AlaGlyThrAspProThrAlaAspProLeuProProThrSerValSerSerGlyGlyAsp 421
QY 1120 ----- 1120
Db 422 AlaProGlyThrValAlaProThrAlaValGluGlySerValAlaAlaGlyValGlyThr 441
QY 1120 ----- 1120
Db 442 AlaGluAspValAlaAlaAlaThrProGlyAlaThrAlaAlaAspValAlaValAspThr 461
QY 1120 ----- 1120
Db 462 AlaGlyAlaThrAspGlyAspAlaValGlyProThrAlaAlaAlaThrAlaGluSerIle 481
QY 1120 ----- 1120
Db 482 AlaAspProThrAlaGlyAlaThrAspGlyAspAlaValGlyAlaThrAlaGluSerIle 501
QY 1120 ----- 1120
Db 502 AlaAspProThrAlaGlyAlaThrAspGlyAspAlaValGlyProThrAlaAlaAlaThr 521
QY 1120 ----- 1120
Db 522 AlaGluSerIleAlaAspProThrAlaGlyAlaThrAlaValSerSerGlySerAlaThr 541
QY 1120 ----- 1120
Db 542 AlaGlyAlaThrAlaGluProLeuLeuValIysArgGlnAlaProGluAlaGluPro 561
QY 1121 GAAACTGCCAGATTACAGATATGGCCACTTTCAAGCCACCATCACAAATTGTAGAGGA 1180
Db 562 ThrGlyCysValLeuTyArgTyArgTyGlyThrPheSerThrGluLeuAsnIleValGlnGly 581
QY 1181 ATCTTAGAGGTTAACATCATCCAGATGACAGAGCTCTGTATGCGGTGCCATGCCCTGAA 1240
Db 582 IleGluSerValaIleValGln-----ValValProAlaAlaProGluGlySer 598
QY 1241 AGCTCCCTAATAGACTTTCGTGACCTGCAAGAGGAGATTCCACGAGGCTCTGTACC 1300
Db 599 GlyAsnSerValGluLeuThrValThrCysGluGlySerLeuProGluGluValCysThr 618
QY 1301 ATCATTTTGACCCCACTGCGAGATCATCCAGACACACAGCTCTGACGCTGTGGATGTG 1360
Db 619 ValValaAspAlaGluCysArgThrAlaGlnMetGlnThrCysSerAlaValAlaPro 638
QY 1361 GATGATGTGCTGTGCTGTGAGACGACCTTCAATGGGTCTGGGACGTACTGTGTG 1420
Db 639 AlaProGlyCysGlnLeuValLeuArgGlnAspPheAsnGlnSerGlyLeuTyCysLeu 658
QY 1421 AACCTCACCTGGGGATGACACAAACCTGGCTCTCACAGACACCTGATT----- 1471
Db 659 AsnValSerLeuAlaAsnGlyAsnGlyLeuAlaValAlaSerThrHisValaValGly 678
QY 1472 TCTGTTCTCTGACAGA 1486

```

```

Db 679 SerIleProSerArg 683
RESULT 5
PM17 HUMAN
ID PM17 HUMAN STANDARD; PRT; 661 AA.
AC P40567; Q12763; Q14448; Q14817; Q16565;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Melanocyte protein Pmel 17 precursor (Melanocyte lineage-specific
DE antigen gp100) (Melanoma-associated ME20 antigen) (ME20M/ME20S)
DE (ME20-M/ME20-S) (95 kDa melanocyte-specific secreted glycoprotein).
GN SILV OR Pmel17 OR D12853E.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RX MEDLINE=92021023; PubMed=1924386;
RA Kwon B.S., Chintamaneni C., Kozak C.A., Copeland N.G.,
RA Gilbert D.J., Jenkins N.A., Barton D., Francke U., Kobayashi Y.,
RA Kim K.-K.;
RT "A melanocyte-specific gene, Pmel 17, maps near the silver coat color
RT locus on mouse chromosome 10 and is in a syntenic region on human
RT chromosome 12."
RL Proc. Natl. Acad. Sci. U.S.A. 88:9228-9232 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94327568; PubMed=7519602;
RA Adema G.J., de Boer A.J., Vogel A.M., Loenen W.A., Figdor C.G.;
RT "Molecular characterization of the melanocyte lineage-specific
RT antigen gp100."
RL J. Biol. Chem. 269:20126-20133 (1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96154052; PubMed=8592076;
RA Bailin T., Lee S.-T., Spritz R.A.;
RT "Genomic organization and sequence of D12853E (Pmel 17), the human
RT homologue of the mouse silver (si) locus."
RL J. Invest. Dermatol. 106:24-27 (1996).
RN [4]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-53.
RX MEDLINE=94235165; PubMed=8179825;
RA Mareh G.A., Marken J.S., Neubauer M., Aruffo A., Hellstrom I.,
RA Hellstrom K.E., Marquardt H.;
RT "Cloning and expression of the gene for the melanoma-associated ME20
RT antigen."
RL DNA Cell Biol. 13:87-95 (1994).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=96314705; PubMed=8739560;
RA Kim K.-X., Youn B.S., Heng H.H., Shi X.-M., Tsui L.-C., Lee Z.H.,
RA Pickard R.T., Kwon B.S.;
RT "Genomic organization and FISH mapping of human Pmel 17, the putative
RT silver locus."
RL Pigment Cell Res. 9:42-48 (1996).
RN [6]
RP SEQUENCE FROM N.A.
RX Vogel A.;
RL Submitted (NOV-1990) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,

```





Alignment Scores: 4.22e-29 Length: 491  
 Pred. No.: 436.50 Matches: 124  
 Score: 40.22% Conservative: 63  
 Percent Similarity: 26.67% Mismatches: 123  
 Best Local Similarity: 9.00% Indels: 155  
 Query Match: 1 Gaps: 14  
 DB:

US-10-039-272-1 (1-2661) x PM17\_BOVIN (1-491)

Qy 603 CAGTATTCAGAAATGGGACGATGTTCTGAGAGAGTTTCTGGAACACAGCCCAATGTG 662  
 Db 1 GlnTyrPrgInValLeuGlyGlyProValSerGlyLeuSerIleGlyThrAspLysAla 20  
 Qy 663 ACACCTGGGCGCTCAACTCATGGAAGTGAAGTCTCTACAGAACATGGA---CGGGCATAT 719  
 Db 21 MetLeuGlyThrTyrAsnMetGluValThrValTyrHisArgGlySerGlnSerTyr 40  
 Qy 720 GTTCCCATCGACAGTGAAGATGTGTACGTGCTAACAGATCAGATTCCTGTGTGTGTG 779  
 Db 41 ValProLeuAlaHisSerSerAlaPheThrIleThrAspGlnValProPheSerVal 60  
 Qy 780 ACTATGTTCCAG-----AAGAACGATCGAAATTCATCCGACGAAACC 821  
 Db 61 SerValSerGlnLeuGlnAlaLeuAspGlyArgAsnLysArg--PheLeuArgLysGln- 79  
 Qy 822 TTCCCAAGATCTCCCATATGTTGATGTCCTGATTCATGATCCTAGCAGCCTCTCTCA 881  
 Db 80 -----ProLeuThrPheAlaLeuGlnLeuHisAspProSerGlyTyrLeuA 95  
 Qy 882 ATTATTCACATTAACATGACGAGCTTCGGGATAATACCTGCTGTGTGTTTCCA 941  
 Db 95 LaGlyAlaAspLeuSerTyrThrTyrPheGlyAspSerThrGlyThrLeuIleSerA 115  
 Qy 942 CCAATCATCTGTAATCAGATGATGCTCAATGGAACCTTCAGCCTTAACCTCACTG 1001  
 Db 115 rgAlaLeuThrValThrHisThrTyrLeuGluSerGlyProValThrAlaGlnValVal 135  
 Qy 1002 TGAAGCTGACGACCA-----GGACCTGTCGGCCAGCCGACCCACACCA 1049  
 Db 135 euGlnAlaAlaIleProLeuThrSerCysGlySerSerProValProGlyThrThrAspA 155  
 Qy 1050 GA----- 1051  
 Db 155 rgHisValThrThrAlaGluAlaProGlyThrThrAlaGlyGlnValProThrThrGluV 175  
 Qy 1051 ----- 1051  
 Db 175 alMetGlyThrThrProGlyGlnValProThrAlaGluAlaProGlyThrThrValGlyT 195  
 Qy 1051 ----- 1051  
 Db 195 rpValProThrThrGluAspValGlyThrThrProGluGlnValAlaThrSerLysVal 215  
 Qy 1051 ----- 1051  
 Db 215 euSerThrThrProValGluMetProThrAlaLysAlaThrGlyArgThrProGluValS 235  
 Qy 1052 -----CCTTCAAAACCCACCCCTCTTTAGGA----- 1078  
 Db 235 erThrThrGluProSerGlyThrThrValThrGlnGlyThrThrProGluLeuValGluT 255  
 Qy 1079 -----CCTGCTGTGTGACAAC-----C 1094  
 Db 255 hrThrAlaGlyGluValSerThrProGluProAlaGlySerAsnThrSerSerPheMetP 275  
 Qy 1095 CCTGGAG-----CTGAGTAGGATTCCTGATGAA----- 1123  
 Db 275 roThrGluGlyThrAlaGlySerLeuSerProLeuProAspAspThrAlaThrLeuValL 295  
 Qy 1124 -----AACTGCCAGATTAACAGATATGGCCACTTTCAAGCCA 1160  
 Db 1124 ----- 1124

RESULT 7

PTN4\_HUMAN  
 ID PTN4\_HUMAN STANDARD; PRT; 926 AA.  
 AC P29074;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Protein tyrosine phosphatase, non-receptor type 4 (EC 3.1.3.48)  
 DE (protein-tyrosine phosphatase MEG) (PTPase-MEG) (MEG).  
 GN PTPN4.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP MEDLINE=91288564; PubMed=1648233;  
 RA Gu M., York J.D., Warshawsky I., Majerus P.W.;  
 RT "Identification, cloning, and expression of a cytosolic megakaryocyte  
 protein-tyrosine-phosphatase with sequence homology to cytoskeletal  
 protein 4.1."; Proc. Natl. Acad. Sci. U.S.A. 88:5867-5871 (1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Colon;  
 RX MEDLINE=22398257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S., Wang J., Hong F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Scaplenko M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abranson R.D., Mullaby S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

	201	TCTTCTGATGAATAAGTACACTGGAAATGAAGAACAATCTAACCAGTGTTGGAAACGCCGGGAGACATG	260
Db	224	AlaArgAspGlnSerAsn---AsnGluIleMetIleGlyValMetSerGlyGlyLeu	247
	261	AGTGTAAGAAAACCTCCTCGGAAGGACGCCGTGTGCAGCCGGTCTGACCAGTGTACTCACCA	320
Db	243	IleTyLysAsn-----ArgValArgMetAsnThrPhePro	254
	321	GCCCTCGTGGGCTCAAATAATAACATTGGGTGAACCTGATATTCCTTAGATCCCCAAAAG	380
Db	255	TrpLeuLysIleValLysIleSerPheLysCysLysGlnPhePheIleGlnLeuArgLys	274
	381	GAAGATGCCAAATGGC-----AACATGTCTCATGAGAAAGAACTGC	419
Db	275	GluLeuHisGluSerArgGluThrLeuLeuGlyPheAsnMetValAsnTyrArgAlaCys	294
	420	AGAAATGAGGCTGGTTTTATCTCTGTCATCATATGTTTACAACTGGACACATGGTCAGAG	479
Db	295	LysAsn-----LeuTrpLysAlaCysValGlu	303
	480	GACATGACGGGAAAAATGGCACCGGCCAAAGCCCATCATACGTCTTCCTCTCATGGAAA	539
Db	304	-----HisHisThrPhePheArgLeuAspArg	312
	540	CCTTTTCTCTCACCCC CGGATGGAGACATGGAAATTTTCATCTACGTCTTCCACACACTT	599
Db	313	ProLeuProProGlnLys-----AsnPhePheAlaHisTyrPheThrLeu	327
	600	GGTCAGTATTTCCAGAAATTTGGACGATGTTTCAGTG-----	635
Db	328	GlySerLysPheArgTyrCysGlyArgThrGluValGlnSerValGlnTyrGlyLysGlu	347
	636	-----AGAGTTCTGTGAACACACAGCCAATGTGACACTGGCCCTCAACTC	680
Db	348	LysAlaAsnLysAspArgValPheAlaArgSerProSerLysProLeuAlaArgLysLeu	367
	681	ATGGAAGTGACTGTCTACAGAAGACATGGACGGGCATATGTTCCATCGGCACAAAGTGAAA	740
Db	368	MetAspTrpGluValValSerArgAsn-----SerileSer	379
	741	GATCTGTACGTGTTACAGATCATCTCTGTGTTGTGATATGTTTCCAGAAAGACGAT	800
Db	380	AspAspArgLeuGluThrGlnSerLeuProSerArgSerProProGlyThrProAsnHis	399
	801	CGAAATTCATCCGA-----CGAAACCTTCCCAAGATCTCCCATATGAT	844
Db	400	ArgAsnSerThrPheThrGlnGluGlyThyArgLeuArgProSerValGlyHisLeu	419
	845	TTTGATGTCTGATTTCAT---GATCTAGCCAC--TTCTCATATTATTCTACCATTAAC	898
Db	420	ValAspHisMetValHisThrSerProSerGluValPheValAsnGlnArg-----	436
	899	TACAACTGAGCTTCGGGATAATATCTGGCCTGTTTGTTTCCACCAATCATACTGTGAAT	958
Db	437	-----SerProSerSerThrGlnAla	443
	959	CACACGTATGTGCTCAATGGAACCTTCAGCCCTTAACCTCACTGTGMAAGCTGAGCACCA	1018
Db	444	AsnSerIleValLeuGluSerSerProSer-----GlnGluThrPro	457
	1019	GGACCTTGTCCGCCACCGCCACACCAACCCAGA-----	1051
Db	458	GlyAspGlyLysProProAlaLeuProProLysGlnSerLysLysAsnSerTrpAsnGln	477
	1051	-----	1051
	478	IleHisTyrSerHisSerGlnGlnAspLeuGluSerHisIleAsnGluThrPheAspIle	497
Db	1052	-----CCTTCAAACCCACCCCTCTTTAGAACCTGTGGTGCACAAACCCCTGGAG	1102
	498	ProSerSerProGluLysProThrProAsnGlyGlyIleProHisAspAsnValLeu	517







```

581 sAspSerProArg-----ProArg 587
371 ATGCCAAAAGGAGATGCCAATGGCAATAGTCTATGAGAAAGTGC---AGAATGA 427
587 gCysGlyGlyLeuGluCysLeuGlyProSerIleHisIleAlaAsnCysSerArgAsnG1 607
428 GCGTGGTTTATCTGCTGATCCATATGTTTACAACTGGACAGCATGTGTCAGAGACAGTGA 487
607 yAla-----TrpThrAlaTrpSerSerTrpAlaG1 617
488 CGGGGAAATGGCACCGGC-----CAAAGCCATCATAACTGCTT 526
617 nCysSerThrSerCysGlyIleGlyPheGlnValArgGlnArgSerCysSerAsn----- 635
527 CCCTGATGGGAAACCTTTCTCCACACCCCGGATGAGAGATGGAATTCATCTACGT 586
636 -----ProAlaProArgHisGlyGly----- 642
587 CTTCACACACTGGTTCAGTATTTCCAGAA-----ATTGGGACGATGTTCACTGAGAGTTT 642
643 -----ArgIleCysValGlyLysSerArgGluGluAArgPh 654
643 CTGT---GAACACAGCAATGTGACACT-----TGGCCCTCAACTCATGGA---AG 687
654 eCysAsnGluAsnThrProCysProValProIlePheTrpAlaSerTrpGlySerTrpSe 674
688 TGACTGTCTACAGAAGACATGAGCGGCATATGTTCCATCGCACAGTGAAGATGTGT 747
674 rLysCysSerAsnAsnCysGlyGlyValGlnSerArgArgSer-----CysG1 692
748 ACCTGGTAAACAGATCAGATTCCTGTTGTGACTAT-----GT 786
692 uAsnGlyAsnSerCysProGlyCysGlyValGluPheLysThrCysAsnProGluAlaCy 712
787 TCCAGAAGACGATCGAAATTCATCCGACGAAACCTTCCCAAGAT-----CTCCCAT 840
712 sProGlu-----ValArgAsnThrProTrpThrProTrpLeuProVa 727
841 TATGTTTGTGAT-----GTCT 855
727 IAsnValThrGlnGlyGlyAlaArgGlnGluArgPheArgPheThrCysArgAlaPr 747
856 GATTCATGATCTGACCTGCTCTCAATTAATTTCT-----ACCAATTAACACAGTGGAG 909
747 oLeuProAspPro---HisGlyLeuGlnPheGlyLysArgArgThrGluThrArgThrCy 766
910 CTTGCGGGATAATCTGGCTCTTT----- 934
766 sProAlaAspGlyThrGlyAlaCysAspThrAspAlaLeuValGluAspLeuLeuArgSe 786
935 -----GTTTCCCAATCATCTGATGATCATCAGTATGTGCTCAATGGAACTTCAGC-- 988
786 rGlySerThrSerProHisThrLeuAsnGlyGlyTrpAlaThrTrpGlyProTrpSerSe 806
989 -----CTTAACCTCACTGTGAAGCTGCGACGACGACGACGACCTTGTCC 1029
806 rCysSerArgAspCysGluLeuGlyPheArgValArgLysArgThr-----CysTh 823
1030 GCCACCGCCACACACCCAGACCTTCAAAACCCACCTCTTTTAGGACCTCTGCTGGA 1089
823 rAsnProGluPro-----ArgAsnGlyGlyLeuProCysValGlyAspAlaAlaG1 840
1090 C-----AACCCCTGGAGCTGAGTAGGATCTCGATGAAACTGCCAGATTAA 1137
840 uTyrglnAspCysAsnPro-----GlnAlaCysProValAr 852
1138 CAGATAT-----GGCCA 1149
852 gglyAlaTrpSerCysTrpThrAlaTrpSerGlnCysSerAlaSerCysGlyGlyHI 872
1150 CTTTCAAGCCACCATCATCAATTTAGAGGAATCTTAGAGGTTAACATCATCCAGATGAC 1209
872 sTyrglnArgThrArgSerCysThrSer----- 881
1210 AGAGTCTCTGATCCCGTGCCTGAGAGCTCCCTAAATAGACTTTTCTGCTGACCTG 1269
882 -----ProAlaProSerProGlyGluAspIleCysLeuGlyLeuHisThrG1 897
1270 CCNAGGGAGCATTCGCCAGGAGTCTGTACCATCATTTCTGACCCACCTGCGAGATCAC 1329
897 uGluAlaLeuCysSerThrGlnAlaCys-Pro----- 907
1330 CCAGAACACAGTCTGCAGCCCTGTGGATG---TGGATGAGATGTCTGCTGACTGTGAG 1386
908 -----GluGlyTrpSerLeuTrpSerGluTrpGlyValCysThr-----G 921
1387 ACNAGCCTTCAATGGGTCTGGGAGTACTGTGTGAA-----CCTACCCCTGGG--- 1434
921 luAspGlyAlaGlnSerArgSerArgSer-CysGluGluLeuLeuProGlyProGlyAla 940
1435 -----GGATGACACAGCCTGCTCTCACGAGCACCTGATTTCTCTCTGACAGAGA 1488
941 CysValGly-AsnSerSerGlnSerArgProCysProTyrrSerGluIleProValIleLe 960
1489 CCCAGCCTCGCTTTAAGGATGGCAACAGTGC-----CTCAT 1527
960 uProAlaSerSerValGluGluThrThrSerCysGlyGlyPheAsnLeuIleHisLeuI1 980
1528 CTCGGTGGTGTCTGGCCATATTTGTCACT---GTGATCTCCCTCTTGGGTGAC----- 1579
980 eValThrGlyValSerCysPheLeuValSerGlyLeuLeuThrLeuAlaValTyLeuSe 1000
1580 -----AAAAACACAGGAATACACCAATAGAAAAATAGTCTGGGAA 1623
1000 rCysGlnHisCysGlnArgGlnSerGlnGluSerThrLeuValHisProAlaThrProAs 1020
1624 TGTGTCTAGAACCAAGCCCTGAGTGTCTTTCTCAACCGTGCAAAAGCGTGTCTTCCC 1683
1020 nHisLeuHisTyrrLysGlyGlyGlyThr----- 1729
1684 GGGAAACACGAGAAAGGATCGCTACTCAAAACCAAGAAATTA 1729
1030 oLysAsnGluLysTyrrThrProMet-----GluPheLys 1041

```

```

RESULT 10
ID - HRG_HUMAN STANDARD; PRT; 525 AA.
AC P04196;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Histidine-rich glycoprotein precursor (Histidine-proline rich
DE glycoprotein) (HPRG).
GN HRG.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=86216149; PubMed=3011081;
RA Koide T., Foster D.C., Yoshitake S., Davie E.W.;
RT "Amino acid sequence of human histidine-rich glycoprotein derived
RT from the nucleotide sequence of its cDNA.";
RL Biochemistry 25:2220-2225(1986).
RN [2]
RP SEQUENCE FROM N.A.
RA Wakabayashi S., Takahashi K., Tokunaga F., Koide T.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 214-247 FROM N.A.
RX MEDLINE=94245171; PubMed=8188234;
RA Hennis B.C., Frants R.R., Bakker E., Vossen R.H., van der Poort E.W.,
RA Blonden L.A., Cox S., Khan P.M., Spurr N.K., Kluff C.;
RA "Evidence for the absence of intron H of the histidine-rich
RT

```

RT glycoprotein (HRG) gene: genetic mapping and in situ localization of  
 RT HRG to chromosome 3q28-q29.;  
 RL Genomics 19:195-197(1994).  
 RN [4]

RP SEQUENCE OF 19-27.  
 RC TISSUE=Plasma;  
 RX MEDLINE=93092937; PubMed=1459097;  
 RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,  
 RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,  
 RA Hochstrasser D.F.;  
 RT "Plasma protein map: an update by microsequencing.";  
 RL Electrophoresis 13:707-714(1992).

CC -!- FUNCTION: The physiological function is not yet known. It binds  
 CC heme, dyes and divalent metal ions. It can inhibit rosette  
 CC formation and is known to interact with heparin, thrombospondin,  
 CC and the lysine-binding site of plasminogen. On the basis of its  
 CC homology with HMW kininogen, the His-rich region of this protein  
 CC may mediate the contact activation phase of intrinsic blood  
 CC coagulation cascade.

CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.  
 CC -!- DOMAIN: In addition to having a high His and Pro content, this  
 CC protein has many internal repeats. 12 tandem repetitions of a 5-  
 CC residue sequence (GHPPH, consensus) form a histidine-rich region.  
 CC -!- SIMILARITY: Contains 2 cystatin-like domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M13149; AAA52694.1; -;  
 DR EMBL; AB005803; BAA21613.1; -;  
 DR EMBL; Z17218; CAA78925.1; -;  
 DR PIR; A01287; KGHUGH.  
 DR SWISS-2DPAGE; P04136; HUMAN.  
 DR Genew; HGNC:5181; HRG.  
 DR MIM; 142640; -;

DR InterPro; IPR000010; Cystatin.  
 DR Pfam; PF00031; cystatin; 1.

DR SMART; SM00043; CY; 2.

KW Glycoprotein; Heparin-binding; Repeat; Signal; Polymorphism.

FT CHAIN 1 18  
 FT DOMAIN 19 525 HISTIDINE-RICH GLYCOPROTEIN.  
 FT DOMAIN 19 136 CYSTATIN-LIKE 1.  
 FT DOMAIN 137 254 CYSTATIN-LIKE 2.  
 FT DOMAIN 276 321 PRO-RICH.  
 FT DOMAIN 350 497 BY SIMILARITY.  
 FT DISULFID 24 504 BY SIMILARITY.  
 FT DISULFID 78 89 BY SIMILARITY.  
 FT DISULFID 105 126 BY SIMILARITY.  
 FT DISULFID 203 417 BY SIMILARITY.  
 FT DISULFID 218 241 BY SIMILARITY.  
 FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 344 344 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 345 345 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARIANT 204 204 P -> S (in dbSNP:3181917).  
 FT /FTID=VAR 014528.

SQ SEQUENCE 525 AA; 59578 MW; A2B124D6CE93114F CRC64;

Alignment Scores:

Pred. No.: 0.295 Length: 525  
 Score: 111.00 Matches: 51  
 Percent Similarity: 29.50% Conservative: 8  
 Best Local Similarity: 25.50% Mismatches: 51  
 Query Match: 2.29% Indels: 91  
 DB: 1 Gaps: 13

US-10-039-272-1 (1-2661) x HRG\_HUMAN (1-525)  
 QY 1028 CGCCACCGCCACCA-----CCACCCAGACCTTCAAAACCC 1063  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 297 ProProProProProProProProProProProProProProProPro 316  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 1064 ACCCTTCTTTAGGACCTGCTGGTGACAAACCCCTGGAGCTAGTAGTCTGATGAA 1123  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 317 ProProLeuLeu-----ProMetSerCysSer 325  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 1124 AACTGCCAGATTAAAGATATGCCCA----- 1149  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 326 SerCysGlnHisAlaThrPheGlyThrAsnGlyAlaGlnArgHisSerHisAsnAsn 345  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 1150 -----CTTTCAGCCAC-----CATCAAAATTAGAGGAATCTTAGAGTTAAACAT 1197  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 346 SerSerAspLeuHisProHisLysHisSerHisGluGlnHisProHisGly--His 364  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 1198 CATCCAGATGACAGACGCTCTGATCGCGTGCCATGGCTGAAAGCTCCCTAATAGACTT 1257  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 365 HisProHisAlaHisHisProHisGlu----- 373  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 1258 TGTCTGACCTGCCAAGGAGCATTCACGAGGCTCTGTACCATCATTTCTGACCCAC 1317  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 374 ---HisAspThrHisArgGlnHisProHisGly-----HisHis-----ProHis 387  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 1318 CTGCGAGATCACCCAGACACAGTCGAGCCCTGTGATGTGGATGGATGATGTGTCTGCT 1377  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 388 -----GlyHisHisProHisGlyHisHisProHisGly----- 398  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 1378 GACTGTGAGACGAACCTTCAATGGTCTGGGACCTACTGTGAACTCACCCTGGGGGA 1437  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 399 -----HisHisProHisGly 403  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 1438 TGACAAAGCTGCTCTCACGACACCTGATTTCTTCTCTCACAGACCCAGCCCTC 1497  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 404 ---HisHisProHisCysHisAsp----- 410  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 1498 GCCTTTAAGGATGCCAAACAGTGCCTGATCTCGTGGCTGTGGCCATATTGTGAC 1557  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
 QY 411 ---PheGlnAspTyrGlyProCysAspProProHisAsnGlnGlyHisCysCysHis 429  
 DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 11

PCLO\_MOUSE

ID PCLO\_MOUSE STANDARD; PRT; 5038 AA.

AC Q9QYX7; Q9QYX6; Q9QZJ0;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Piccolo protein (Presynaptic cytomatrix protein) (Aczonin) (Brain-

DE derived HLMN protein).

GN PCLO OR ACZ.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, ALTERNATIVE SPLICING,

RP TISSUE SPECIFICITY, AND INTERACTION WITH PROFILIN.

RC TISSUE=Brain;

RX MEDLINE=99439764; PubMed=10508862;

RA Wang X., Kibschull M., Laue M.M., Lichte B., Petrasch-Parwez E.,

RA Kilimann M.W.;

RT "Aczonin, a 550-kd putative scaffolding protein of presynaptic active

RT zones, shares homology regions with rim and bassoon and binds

RT profilin.";

RL J. Cell Biol. 147:151-162(1999).

RP REVISIONS.

RA Kilimann M.W.;

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 4502-4682 FROM N.A.

TISSUE=Brain;  
Huang W., Jin W., Huang C., Chen B., Zhang J., Ju G.;  
Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
[4]  
INTERACTION WITH RIMS2.  
MEDLINE=22384373; PubMed=12401793;  
Fujimoto K., Shibasaki T., Yokoi N., Kashima Y., Matsumoto M.,  
Sasaki T., Tajima N., Iwanaga T., Seino S.;  
Piccolo, a Ca2+ sensor in pancreatic beta-cells. Involvement of  
CAMP-GEFII, Rim2, piccolo complex in Ca2+-dependent exocytosis.;  
J. Biol. Chem. 277:50497-50502(2002).  
-!- FUNCTION: May act as a scaffolding protein involved in the  
organization of synaptic active zones and in synaptic vesicle  
trafficking.  
-!- SUBUNIT: Interacts with Rabac1/Pral, RIMS2 and profilin.  
-!- SUBCELLULAR LOCATION: Concentrated at the presynaptic side of  
synaptic junctions.  
-!- ALTERNATIVE PRODUCTS:  
Event=Alternative splicing; Named isoforms=2;  
Name=1;  
IsoId=Q9QYX7-1; Sequence=Displayed;  
Name=2;  
IsoId=Q9QYX7-2; Sequence=VSP 003928, VSP 003929;  
TISSUE SPECIFICITY: Highly expressed in brain. Low levels found in  
stomach. Not detected in other tissues analyzed including adrenal  
gland, testis and pancreas.  
-!- DOMAIN: C2 domain 1 is involved in binding calcium and  
phospholipids. Calcium binds with low affinity but with high  
specificity and induces a large conformational change.  
-!- SIMILARITY: Contains 2 C2 domains.  
-!- SIMILARITY: Contains 1 PDZ/DHR domain.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; Y19185; CAB60731.2; -;  
EMBL; Y19186; CAB60732.2; -;  
EMBL; AF181269; AAD55786.2; -;  
HSSP; P04410; 1A25.  
MGI; MGI:1349390; Pclo.  
GO; GO:0045202; C:synaptic junction; IDA.  
GO; GO:0005509; F:calcium ion binding; ISS.  
GO; GO:0005544; F:calcium-dependent phospholipid binding; ISS.  
GO; GO:0005522; F:profilin binding; IDA.  
GO; GO:0019933; P:cAMP-mediated signaling; IDA.  
GO; GO:0007010; P:cytoskeleton organization and biogenesis; IDA.  
GO; GO:0030073; P:insulin secretion; IDA.  
GO; GO:0017157; P:regulation of exocytosis; IDA.  
GO; GO:0016800; P:synaptic vesicle targeting; NAS.  
InterPro; IPR000008; C2.  
InterPro; IPR001478; PDZ.  
InterPro; IPR008899; Znf\_piccolo.  
Pfam; PF00168; C2; 2.  
Pfam; PF00595; PDZ; 1.  
Pfam; PF05715; Zf\_piccolo; 2.  
SMART; SM00239; C2; 2.  
SMART; SM00228; PDZ; 1.  
PROSITE; PS00499; C2\_DOMAIN\_1; 1.  
PROSITE; PS00004; C2\_DOMAIN\_2; 2.  
PROSITE; PS0106; PDZ; 1.  
Calcium/phospholipid binding; Metal-binding; Zinc; Zinc-finger;  
Repeat; Alternative splicing.  
DOMAIN 371 470 10 X 10 AA TANDEM APPROXIMATE REPEATS OF  
P-A-K-P-Q-P-Q-P-X.  
C4-TYPE (POTENTIAL).  
C4-TYPE (POTENTIAL).  
POLY-POLY.  
PDZ.  
-----  
ZN\_FING 502 526  
ZN\_FING 967 990  
DOMAIN 2305 2329  
PDZ.  
-----  
FT DOMAIN 4607 4705 C2 DOMAIN 1.  
FT DOMAIN 4922 5012 C2 DOMAIN 2.  
FT VARSPLIC 4829 4833 TRPTN -> SKRRK (in isoform 2).  
FT VARSPLIC 4834 5038 /FTId=VSP 003928.  
FT VARSPLIC 4834 5038 Missing (in isoform 2).  
FT SEQUENCE 5038 AA; 547600 MW; DADA460CF3B40888 CRC64;  
SQ  
Alignment Scores:  
Pred. No.: 0.642 Length: 5038  
Score: 111.00 Matches: 23  
Percent Similarity: 60.38% Conservative: 9  
Best Local Similarity: 43.40% Mismatches: 19  
Query Match: 2.29% Indels: 2  
DB: 1 Gaps: 1  
US-10-039-272-1 (1-2661) x PCLO\_MOUSE (1-5038)  
QY 926 GGCTGTTTGTTCACCAATCATCTGTGAATCACACGATGTGTCTCAATGGAACCTTC 985  
Db 2280 GlyLeuProValSerThrHisProSerLysSerHisProPheArgSerSer----- 2297  
QY 986 AGCCTTAACCTCACTGTGAAGCTGCAGCAGCAGGACCTGTGTCGCCACCGCCACCA 1045  
Db 2298 SerLeuAspIleSerAlaGlnProProProProProProProProProProPro 2317  
QY 1046 CCCAGACCTTCAAAACCCACCCCTTCTTAGGACCTGCT 1084  
Db 2318 ProProProProProProProProProProProProProProProProProPro 2330  
RESULT 12  
POLR\_KYMVJ STANDARD; PRT; 1874 AA.  
ID POLR\_KYMVJ  
AC P36304;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE RNA replicase polypeptide (EC 2.7.7.48).  
OS Kennedy yellow mosaic virus (strain Jervis bay) (KYMV).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tymoviridae;  
OC Tymovirus.  
OX NCBI\_TaxID=36392;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=90218040; PubMed=2324710;  
RX Ding S., Keese P., Gibbs A.;  
RT "The nucleotide sequence of the genomic RNA of Kennedy yellow mosaic  
RT tymovirus-Jervis Bay isolate: relationships with potex- and  
RT carlaviruses."; J. Gen. Virol. 71:925-931(1990).  
RL J. Gen. Virol. 71:925-931(1990).  
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
CC (RNA) (N).  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; D00637; BAA00532.1; -;  
PIR; JQ0533; JQ0533.  
DR MGROPS; C21.001; -;  
DR InterPro; IPR008043; Peptidase C21.  
DR InterPro; IPR001788; RNA dep RNAPol2.  
DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro; IPR007094; RNA\_pol\_Pstir.  
DR InterPro; IPR000606; Viral\_helicase.  
DR Pfam; PF05381; Peptidase C21; 1.  
DR Pfam; PF00978; RNA dep RNAPol2; 1.  
DR Pfam; PF01443; Viral helicase; 1.  
KW Transferase; RNA-directed RNA polymerase; Polypeptide; ATP-binding.

QY	1436	GATGACACAGCCTGGCTCTCAGCAGCAC	-----CTGATTTCGTCTCTGACAGA	1486
DB	799	ProProThProLeuProMetProLysAsnAsnCysLeuLeuThraAlaValaProSer	818	
QY	1487	GACCCAGCTCGCTCTTAAGG	-----1507	
DB	819	LeuHisIleAsnProHisArgLeuTprSerLeuGlnGluValLeuProAspSerLeu	838	
QY	1508	ATGCAACACAGTGCCTGATCTCCGTGGCTGTGGCCATATTCTGACTGTGACTCTCC	1567	
DB	839	LeuSerAsnSerGluIleAspSerValGlyMetSerThrAspLeuLeuThraAlaLeuSer	858	
QY	1568	CTCTTG 1573		
DB	859	HisLeu 860		
RESULT 13				
CTD2 HUMAN				
ID	CTD2 HUMAN	STANDARD;	PRT; 1225 AA.	
AC	Q9UQB3; O00379; O15390; O43206; O43840; Q13589; Q9UM66; Q9UPM3;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Catenin delta-2 (Delta-catenin) (Neural plakophilin-related ARM-repeat protein) (NPRAP) (Neurojuncin) (G724).			
DE	CTNND2 OR NPRAP.			
GN	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1), AND INTERACTION WITH PRESENILIN 1.			
RX	MEDLINE=97366296; PubMed=9223106;			
RA	Zhou J., Liyanage U., Medina M., Ho C., Simmons A.D., Lovett M., Kosik K.S.;			
RT	"Presenilin 1 interaction in the brain with a novel member of the Armadillo family.;"			
RL	NeuroReport 8:2085-2090 (1997).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM 1), AND FUNCTION.			
RC	TISSUE=Brain;			
RX	MEDLINE=99268158; PubMed=9971746;			
RA	Lu Q., Paredes M., Medina M., Zhou J., Cavallo R., Peifer M., Orecchio L., Kosik K.S.;			
RT	"Delta-catenin, an adhesive junction-associated protein which promotes cell scattering.;"			
RL	J. Cell Biol. 144:519-532 (1999).			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORM 1), AND INTERACTION WITH PRESENILIN 1.			
RC	TISSUE=Fetal brain;			
RX	MEDLINE=99223289; PubMed=10208590;			
RA	Tanahashi H., Tabira T.;			
RT	"Isolation of human delta-catenin and its binding specificity with presenilin 1.;"			
RL	NeuroReport 10:563-568 (1999).			
RN	[4]			
RP	SEQUENCE OF 186-1225 FROM N.A. (ISOFORM 1), AND INTERACTION WITH PRESENILIN 1.			
RC	TISSUE=Brain;			
RX	MEDLINE=99155075; PubMed=10037471;			
RA	Levesque G., Yu G., Nishimura M., Zhang D.M., Levesque L., Yu H., Xu D., Liang Y., Rogawa E., Ikeda M., Duchie M., Murgolo N., Wang L., Vander-Ver P., Bayne M.L., Strader C.D., Rommens J.M., Fraser P.E., St George-Hyslop P.H.;			
RT	"Presenilins interact with armadillo proteins including neural-specific plakophilin-related protein and beta-catenin.;"			
RL	J. Neurochem. 72:999-1008 (1999).			
RN	[5]			
RP	SEQUENCE OF 350-1225 FROM N.A. (ISOFORM 1), AND ALTERNATIVE SPLICING.			
RC	TISSUE=Fetal brain;			
RX	MEDLINE=98002299; PubMed=9342840;			
RA	Paffenholz R., Franke W.W.;			
RT	"Identification and localization of a neural-specific armadillo protein that interacts with presenilin 1.;"			
RL	J. Neurochem. 72:999-1008 (1999).			
RN	[6]			
RP	SEQUENCE OF 350-1225 FROM N.A. (ISOFORM 1), AND ALTERNATIVE SPLICING.			
RC	TISSUE=Fetal brain;			
RX	MEDLINE=98002299; PubMed=9342840;			
RA	Paffenholz R., Franke W.W.;			
RT	"Identification and localization of a neural-specific armadillo protein that interacts with presenilin 1.;"			
RL	J. Neurochem. 72:999-1008 (1999).			
RN	[7]			
RP	SEQUENCE OF 350-1225 FROM N.A. (ISOFORM 1), AND ALTERNATIVE SPLICING.			
RC	TISSUE=Fetal brain;			
RX	MEDLINE=98002299; PubMed=9342840;			
RA	Paffenholz R., Franke W.W.;			
RT	"Identification and localization of a neural-specific armadillo protein that interacts with presenilin 1.;"			
RL	J. Neurochem. 72:999-1008 (1999).			
RN	[8]			
RP	SEQUENCE OF 350-1225 FROM N.A. (ISOFORM 1), AND ALTERNATIVE SPLICING.			
RC	TISSUE=Fetal brain;			
RX	MEDLINE=98002299; PubMed=9342840;			
RA	Paffenholz R., Franke W.W.;			
RT	"Identification and localization of a neural-specific armadillo protein that interacts with presenilin 1.;"			
RL	J. Neurochem. 72:999-1008 (1999).			
RN	[9]			
RP	SEQUENCE OF 350-1225 FROM N.A. (ISOFORM 1), AND ALTERNATIVE SPLICING.			
RC	TISSUE=Fetal brain;			
RX	MEDLINE=98002299; PubMed=9342840;			
RA	Paffenholz R., Franke W.W.;			
RT	"Identification and localization of a neural-specific armadillo protein that interacts with presenilin 1.;"			
RL	J. Neurochem. 72:999-1008 (1999).			
RN	[10]			
RP	SEQUENCE OF 350-1225 FROM N.A. (ISOFORM 1), AND ALTERNATIVE SPLICING.			
RC	TISSUE=Fetal brain;			
RX	MEDLINE=98002299; PubMed=9342840;			

RT plakoglobin/armadillo multigene family.";

RL Differentiation 61:293-304(1997).

RM [6]

RN SEQUENCE OF 543-656 FROM N.A.

RA Wilson R., Smith A., Elliott G., Kramer J., Latreille P., Keppler D.;

RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.

RN [7]

RN SEQUENCE OF 865-1225 FROM N.A.

RP MEDLINE=97202103; PubMed=9049630;

RA Simmons A.D., Overhauser J., Lovett M.;

RT "Isolation of cDNAs from the Cri-du-chat critical region by direct

RT screening of a chromosome 5-specific cDNA library.";

RL Genome Res. 7:118-127(1997).

RN [8]

RN SEQUENCE OF 873-1225 FROM N.A.

RP TISSUE=Brain;

RA Yu W., Sarginson J., Gibbs R.A.;

RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: May be involved in neuronal cell adhesion and tissue

CC morphogenesis and integrity by regulating adhesion molecules.

CC Induces cell motility and cell scattering in response to

CC hepatocyte growth factor treatment (in vitro).

CC -!- SUBUNIT: BINDS TO B-CADHERIN AT A JUNCTIONAL SITE WITHIN THE

CC CYTOPLASMIC DOMAIN (BY SIMILARITY). BINDS TO PRESENILIN-1.

CC -!- SUBCELLULAR LOCATION: Adherens junction (By similarity).

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Comment=Additional isoforms seem to exist;

CC Name=1; Synonyms=NPRAPa;

CC IsoId=Q9UBQ3-1; Sequence=Displayed;

CC Name=2; Synonyms=NPRAPc;

CC IsoId=Q9UBQ3-2; Sequence=VSP\_006746;

CC -!- TISSUE SPECIFICITY: Predominantly expressed in brain.

CC -!- SIMILARITY: Contains 9 ARM repeats.

CC -!- SIMILARITY: Belongs to the beta-catenin family.

CC -!- CAUTION: REF.8 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO FRAMESHIFTS

CC IN POSITION 746 TO 913.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

DR EMBL; U96136; AAC63103.1; -

DR EMBL; U52828; AAB96357.1; -

DR EMBL; AB013805; BAA36163.1; -

DR EMBL; U81004; AAD00453.1; -

DR EMBL; U72665; AAB68599.1; -

DR EMBL; U52351; AAB97957.1; -

DR EMBL; AC003089; AAB83940.1; -

DR EMBL; AF035302; AAB88185.1; ALT\_FRAME.

DR Genew; HGNC:2516; C1ND2.

DR MTM; G04275; -

DR GO; GO:0007158; P:neuronal cell adhesion; NAS.

DR InterPro; IPR008938; ARM.

DR InterPro; IPR000225; Armadillo.

DR Pfam; PF00514; Armadillo\_seg; 5.

DR SMART; SM00185; ARM; 7.

DR PROSITE; PS01076; ARM\_REPEAT; 3.

KW Cell adhesion; Cytoskeleton; Repeat; Coiled coil; Alternative splicing.

KW Developmental protein; Repeat; COILED COIL (POTENTIAL).

FT DOMAIN 49 84

FT DOMAIN 216 226

FT DOMAIN 811 817

FT REPEAT 394 438

FT REPEAT 540 579

FT REPEAT 582 621

FT REPEAT 626 666

FT REPEAT 682 724

FT REPEAT 728 773

FT REPEAT 835 875 ARM 7.

FT REPEAT 882 921 ARM 8.

FT REPEAT 975 1018 ARM 9.

FT VARSPLIC 823 880 Missing (in isoform 2).

FT /FTId=VSP\_006746.

FT CONFLICT 426 426 V -> D (IN REF. 3, 4 AND 5).

FT CONFLICT 443 443 V -> D (IN REF. 3, 4 AND 5).

FT CONFLICT 725 725 P -> A (IN REF. 3, 4 AND 5).

FT CONFLICT 894 894 R -> L (IN REF. 3, 4 AND 5).

FT CONFLICT 909 909 A -> V (IN REF. 3, 4 AND 5).

FT CONFLICT 980 980 A -> G (IN REF. 5).

FT CONFLICT 1102 1102 A -> G (IN REF. 4).

SQ SEQUENCE 1225 AA; 132664 MW; 8B676CFD5AFA4E65 CRC64;

Alignment Scores:

Pred. No.: 0.48 Length: 1225

Score: 110.00 Matches: 87

Percent Similarity: 32.43% Conservative: 44

Best Local Similarity: 21.53% Mismatches: 132

Query Match: 2.27% Indels: 141

DB: 1 Gaps: 19

US-10-039-272-1 (1-2661) x CTD2\_HUMAN (1-1225)

Qy 520 ACCTCTCCCTGATGGAAACCTTTCTCACCCCGGATGGAGAGATGGAATTTC 579

Db 28 ThrSerSerLeuSerProGly-----LeuAsnThrSerAsnGlyAspGlySerGluThr 45

Qy 580 TCTACGCTTCCACACACTTCGTCTCAGTATT----- 609

Db 46 GluThrThrSerAlaLeuLeuAlaSerValLysGluGlnGluLeuGlnPheGluArgLeu 65

Qy 609 ----- 609

Db 66 ThrArgGluLeuGluAlaGluArgGlnIleValAlaSerGlnLeuGluArgCysLysLeu 85

Qy 610 ---TCCAGAAATGGACAGATGTTCACTGAGAGTTCTGTGAACACAGCAATGTCACAC 666

Db 86 GlySer-GluThrGlySerMetSerSerSerAlaGluGluGln-----PheG1 103

Qy 667 TTGGCCCTCACTCATGGA---AGTGACTGTCTACAGAGACAGATGACGGGATGTTTC 723

Db 103 nTrpGlnSerGlnAspGlyGlnLysAspIleGluAspGluLeuThrThrGlyLeu----- 121

Qy 724 CCATCGCACAACTGAGATGTCGTCGTGAACAGATCAGATTCCTGTTGTGTGACTA 783

Db 122 ----GluLeuValAspSerCysIleArg---SerLeuGlnGluSerGlyLeuLeuAsp-- 138

Qy 784 TGTTCAGAGAACGATCGAAATTCATCCGACGAAACCTTCCCAAAGATCTCCCAATTAT 843

Db 139 ----ProGlnAspTyrSerThrGlyGluArgProSerLeuLeuSerGlnSerAlaLeuG1 157

Qy 844 GTTTGATGCTGATTTCATGATCTAGCCACTTCCTCAATTATTTCACCAATTAACCTACAA 903

Db 157 n-----LeuAsnSerLysProGlu-----GlySerPheGlnTyrPr 169

Qy 904 GTGGAGCTTCGGGATAATACTGCGCTGTTGTTTCCCAATCATCATCTGTCGATCACAC 963

Db 169 oAlaSerTyrHisSerAsnGlnThrLeu-----AlaLeuGlyGluThrThrProSerG1 187

Qy 964 GTATGTGCTCAATGGAAC-----TTCAGCCCTTAACCTCAC 999

Db 187 nLeuProAlaAArgGlyThrGlnAlaArgAlaThrGlyGlnSerPheSerGlnGlyThrTh 207

Qy 1000 TGTGAAGCT-----GCAGCACGAGGACCTTGTTCGCCACCGCCACACACCCAG 1050

Db 207 rSerArgAlaGlyHisLeuAlaGlyProGluProAlaProProProProProAr 227

Qy 1051 A----- 1051

Db 227 gGluProPheAlaProSerLeuGlySerAlaPheHisLeuProAspAlaProAlaAl 247

Qy 1052 -----CCTTCAAAACCC-----AC 1065

```

Db      247  aalaAlaAlaLeuTyTySerSerThrThrLeuProAlaProProArgGlyGlySe 267
QY      1066  CCCITCTTTAGACCTGCTGTGTGACACACCCCTGGAGCTGAGTAGG-----ATTCC 1116
Db      267  iProLeuAlaAlaProGlnGlyGlySerProThrThrLysLeuGlnArgGlyGlySerAlaPr 287
QY      1117  TGATGAAACTGCCAGATTAAACAGATATGGCCACTTTCACGCCACCATCAAAATTGTAGA 1176
Db      287  oGlu-----GlyAlaThrTyAlaAlaProAr 296
QY      1177  GGGAACTTAGAGTTAATCATCATCCAGATGACACGCTCTGTATGCCGGTGCATCGCC 1236
Db      296  gGly-----SerSerProLysGlnSerProSerArgLeuAlaLysSerTySe 312
QY      1237  TGAAGCTCCCTAATAGACTTTCGTGACCTGCCAGGAGCATTCACCGAGGCTG 1296
Db      312  rThrSerSerProLysGlnValValSerSerAlaGlyLeuSerProLysArgVal-- 331
QY      1297  TACCATCATTTCTGACCCCTCGAGATCAACCCAGACACACATCTCGACCCCTGTGGA 1356
Db      332  -----ThrSerPro----- 335
QY      1357  TGTGGATGAGATGTCTGTGCTGAGTGTGAGACGAACTTCAATGGTCTGGGACCTACTG 1416
Db      336  -----ThrValGlnSerThrIleSerSerSerProLysHisG1 348
QY      1417  TGTGAACCTCACCTGGGGATGACACACGCTGGCTCTCAGAGACCCCTGATTTCTGT 1476
Db      348  nLeuSerSerThrIleGlyThrTyAlaThrLeuSerProThrLysArgLeuValHisAl 368
QY      1477  TCCTGACAGA 1486
Db      368  aSerGluGln 371

RESULT 14
MUSE_HUMAN
ID   MUSEB_HUMAN          STANDARD;          PRT;          5703 AA.
AC   Q9HC84; O00447; O14985; O15494; O95291; O95451; Q14881;
CD   Q95552; Q9UE28;
DT   28-FEB-2003 (Rel. 41, Created)
DE   Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High
GN   molecular weight salivary mucin MG1) (Sublingual gland mucin).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN   [1]
RP   SEQUENCE OF 1-1594 FROM N.A.
RA   Chen Y., Di Y.P., Wu R.;
RT   "Molecular cloning of the amino-terminal and 5'-flanking region of the
RL   human MUC5B mucin gene.";
RN   [2]
RP   SEQUENCE OF 1-1325 FROM N.A.
RA   Desseyn J.-L., Buisine M.P., Porchet N., Aubert J.-P., Laine A.;
RT   "Genomic organization of the human mucin gene MUC5B: cDNA and genomic
RN   sequences upstream of the large central exon.";
RN   [3]
RP   SEQUENCE OF 40-1324 FROM N.A.
RA   Desseyn J.-L., Buisine M.P., Porchet N., Aubert J.-P., Laine A.;
RT   "Genomic organization of the human mucin gene MUC5B: cDNA and genomic
RN   sequences upstream of the large central exon.";
RN   [4]
RP   SEQUENCE OF 1326-4895 FROM N.A.

```

```

RC      TISSUE=Placenta;
RX      MEDLINE=97166151; PubMed=9013550;
RA      Laine A.;
RT   "Human mucin gene MUC5B, the 10.7 kb large central exon encodes
RT   various alternate subdomains resulting in a super-repeat. Structural
RL   J. Biol. Chem. 272:3168-3178(1997).
RN   [5]
RP   SEQUENCE OF 4057-4480 FROM N.A.
RC      TISSUE=Salivary gland;
RX      MEDLINE=97292540; PubMed=9147051;
RA      Nielsen P.A., Bennett E.P., Wandall H.H., Therkildsen M.H.,
RA      Hannibal J., Clausen H.;
RT   "Identification of a major human high molecular weight salivary mucin
RT   (MG1) as tracheobronchial mucin MUC5B.";
RL   Glycobiology 7:413-419(1997).
RN   [6]
RP   SEQUENCE OF 4721-5703 FROM N.A.
RC      TISSUE=Gall bladder;
RX      MEDLINE=97293229; PubMed=9164870;
RA      Keates A.C., Nunes D.P., Afzal N.H., Troxler R.F., Offner G.D.;
RT   "Molecular cloning of a major human gall bladder mucin: complete C-
RT   terminal sequence and genomic organization of MUC5B.";
RL   Biochem. J. 324:295-303(1997).
RN   [7]
RP   SEQUENCE OF 4809-5687 FROM N.A.
RC      TISSUE=Sublingual gland;
RX      MEDLINE=96125355; PubMed=8554565;
RA      Troxler R.F., Offner G.D., Zhang F., Iontcheva I., Oppenheim F.G.;
RT   "Molecular cloning of a novel high molecular weight mucin (MG1)
RT   from human sublingual gland.";
RL   Biochem. Biophys. Res. Commun. 217:1112-1119(1995).
RN   [8]
RP   SEQUENCE OF 4859-5703 FROM N.A.
RC      TISSUE=Placenta;
RX      MEDLINE=97347489; PubMed=9201995;
RA      Desseyn J.-L., Aubert J.-P., Porchet N., Laine A.;
RT   "Genomic organization of the 3 region of the human MUC5B mucin.";
RL   J. Biol. Chem. 272:16873-16883(1997).
CC      -!- FUNCTION: Salivary mucin that is thought to contribute to the
CC      lubricating and viscoelastic properties of whole saliva.
CC      -!- SUBCELLULAR LOCATION: Secreted.
CC      -!- TISSUE SPECIFICITY: Expressed mainly in bronchus glands and also
CC      in submaxillary glands, endocervix, gall bladder, and pancreas.
CC      -!- PTM: Highly glycosylated.
CC      -!- SIMILARITY: Contains 1 TIL (Trypsin inhibitory-like) domain.
CC      -!- SIMILARITY: Contains 3 VWFC domains.
CC      -!- SIMILARITY: Contains 4 VWFD domains.
CC      -!- SIMILARITY: Contains 1 C-terminal cysteine knot-like (CTCK) domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AF107890; AAC33673.1; -
DR      EMBL; AF086604; AAC67545.1; -
DR      EMBL; AJ004862; CAA06167.1; -
DR      EMBL; Z72496; CAA96577.1; -
DR      EMBL; X74955; CAA52910.1; -
DR      EMBL; U63836; AAB61398.1; -
DR      EMBL; U78554; AAC51344.1; -
DR      EMBL; U78552; AAC51344.1; JOINED.
DR      EMBL; U78553; AAC51344.1; JOINED.
DR      EMBL; U78551; AAC51343.1; -
DR      EMBL; U95031; AAB65151.1; -
DR      EMBL; Y09788; CAA70926.1; -
DR      Genew; HGNC:7516; MUC5B.
DR      MIM; 600770; -

```





```

QY 907 GAGCTTCGGGGT-----AATAGTGGCTGTTGTTTCCACCAACATCACTGT 954
Db :|||||
QY 4887 gAlaPheGlyGlnPheSerProGlyGluValIleTyrAsnLysThrAspArgAlaGl 4907
QY 955 GAATACACAGTAT---GTCTCAATGGAACTTACGCTTAACCTCACTGCGAAAGCTGC 1011
Db :|||||
QY 4907 yCysHisPheTyrAlaValCysAsnGlnHisCysAspIleAsp-----Ar 4922
QY 1012 AGCACAGGACCTTGTCCGCCACCCACCACCCAGACCTTCAAAACCC-----AC 1065
Db :|||||
QY 4922 gPheGlnGlyAlaCysProThrSerProProVal---SerSerAlaProLeuSerSe 4941
QY 1066 CCCTCTTTAGACCTGTGTGTGTGACACCCCTCGAGCTGAGTAGATTCTCTGAT----- 1120
Db :|||||
QY 4941 rProSerProAlaProGlyCysAspAsnAlaIleProLeuArgGlnValAsnGluThrTr 4961
QY 1121 -----GAAACATGCCAGATTAACAGATATGGCCACTTCAAGCCACCATCAACAATGT 1173
Db :|||||
QY 4961 pThrLeuGluAsnCysThrValAlaArg----- 4970
QY 1174 AGAGGAATCTTAGAGGTTAATCATCATCCAGATGACAGACCTGCTCTGATGCGGTGCCATG 1233
Db :|||||
QY 4971 -----CysValGlyAspAsnArgValValLeuLeuAspProLy 4983
QY 1234 GCCTGAAGCTCCTTAATAGACTTTGTGTGACCTGCCAAGGAGC---ATTCCACCGA 1290
Db :|||||
QY 4983 sProValAlaAsn-----ValThrCysValAsnLysHisLeuProIleLy 4998
QY 1291 GGTCTGTACCATCATTTCTGACCCAC-----TGCAGATCACCCAGAACACAGTCTG 1344
Db :|||||
QY 4998 sVal-----SerAspProSerGlnProCysAspPheHisTyrGluCysGluCy 5014
QY 1345 CAGCCCTGTGGATGTGATGCTGCTGCTG----- 1378
Db :|||||
QY 5014 s-----IleCysSerMetTrpGlySerHisTyrSerThrPh 5027
QY 1379 -----ACTGTGAGACGAACTTCAATGGTCTGGACCTACTGCTGTG----- 1420
Db :|||||
QY 5027 eAspGlyThrSerTyrThrPheArgGlyAsnCysThrTyrValLeuMetArgLulLeHi 5047
QY 1421 -----AAGCTCACCTGGGGGATGACACA----- 1444
Db :|||||
QY 5047 sAlaArgPheGlyAsnLeuSerLeuTyrLeuAspAsnHisTyrCysThrAlaSerAlaTh 5067
QY 1445 -----AGCCTGGC 1452
Db :|||||
QY 5067 xAlaAlaAlaAlaArgCysProArgAlaLeuSerIleHisTyrLysSerMetAspIleVa 5087
QY 1453 TCTCAGACGACCCCTGATTCT-----GTTCTCTGACAGAGACCC 1491
Db :|||||
QY 5087 lLeuThrValThrMetValHisGlyLysGluGluGlyLeuIleLeuPheAspGlnIlePr 5107
QY 1492 AGCCTCGCCTTTAAGATGCGAAACAGTCCCTGATCTCCGTT---GGCTGCTGGCCAT 1548
Db :|||||
QY 5107 oValSerSerGlyPheSerLysAsnGlyValLeuValSerValLeuGlyThrThrWe 5127
QY 1549 ATTCTCTACTGTGATCTCCCTC-----TTGGTGTACAAAA 1584
Db :|||||
QY 5127 tArgValAspIleProAlaLeuGlyValThrValThrPheAsnGlyGlnValPheGlnAl 5147
QY 1585 ACACAGGAATACAAACCAATAGAAATAGTCTCTGG-----GAATGTGT-----CA 1631
Db :|||||
QY 5147 aArgLeuProTyrSerLeuPheHisAsnThrGluGlyGlnCysGlyThrCysThrAs 5167
QY 1632 GAAGCAAGGCGCTGAGTCTCTTCTCAA 1659
Db :|||||
QY 5167 nasnGlnArgAspCysLeuGlnArg 5176

```

RESULT 15

IRF3\_HUMAN

ID IRF3\_HUMAN

AC Q14653

DT 01-NOV-1997

STANDARD; PRT; 427 AA.

01-NOV-1997 (Rel. 35, Last sequence update)  
 16-OCT-2001 (Rel. 40, Last annotation update)  
 Interferon regulatory factor 3 (IRF-3).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 NCBI\_TaxID=9606;  
 [1]  
 SEQUENCE FROM N.A.  
 RC TISSUE=Retina;  
 RX MEDLINE=9610173; PubMed=8524823;  
 RA Au W.-C., Moore P.P.A., Lowther W.W., Juang Y.-T., Pitha P.M.;  
 RT "Identification of a member of the interferon regulatory factor  
 family that binds to the interferon-stimulated response element and  
 activates expression of interferon-induced genes";  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:11657-11661(1995).  
 [2]  
 SEQUENCE OF 323-413 FROM N.A.  
 RX MEDLINE=99020108; PubMed=9803267;  
 RA Bellingham J., Gregory-Evans K., Gregory-Evans C.Y.;  
 RT "Mapping of human interferon regulatory factor 3 (IRF3) to chromosome  
 19q13.3-13.4 by an intragenic polymorphic marker";  
 RL Ann. Hum. Genet. 62:231-234(1998).  
 CC -!- FUNCTION: MAY ACTIVATE TRANSCRIPTION BY COMPLEX FORMATION WITH  
 OTHER TRANSCRIPTIONAL FACTORS, POSSIBLY MEMBERS OF THE STAT  
 FAMILY. BINDS SPECIFICALLY TO THE IFN-STIMULATED RESPONSE ELEMENT  
 (ISRE) BUT NOT TO THE IRF-1 BINDING SITE PRD-1.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- TISSUE SPECIFICITY: Expressed constitutively in a variety of  
 tissues.  
 CC -!- SIMILARITY: Belongs to the IRF family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; Z56281; CAA91227.1; -;  
 CC EMBL; U86636; AAC68818.1; -;  
 CC HSSP; P23906; 2IRF.  
 CC TRANSFAC; T04673; -;  
 CC Genew; HGNC:6118; IRF3.  
 CC MIM; 603734; -;  
 CC GO; GO:0003702; F:RNA polymerase II transcription factor acti. . . ; TAS.  
 CC GO; GO:0003712; F:transcription cofactor activity; TAS.  
 CC GO; GO:0006366; P:transcription from Pol II promoter; TAS.  
 CC InterPro; IPR001346; IRF.  
 CC InterPro; IPR008984; SMAD\_FHA.  
 CC Pfam; PF00605; IRF; 1.  
 CC PRINTS; PR00267; INTFRNREGCT.  
 CC Prodom; PD002355; IRF; 1.  
 CC SMART; SM00348; IRF; 1.  
 CC PROSITE; PS00601; IRF; 1.  
 CC Transcription regulation; DNA-binding; Nuclear protein; Activator;  
 KW Polymorphism.  
 FT DNA BIND 7 107 TRYPTOPHAN PENTAD REPEAT.  
 FT VARIANT 96 96 R -> Q (in dbSNP:968457).  
 FT VARIANT 377 377 E -> K (in dbSNP:1049486).  
 FT VARIANT 427 427 S -> T (in dbSNP:7251).  
 FT SEQUENCE 427 AA; 47219 MW; F536676FA78B0110 CRC64;  
 Alignment Scores:  
 Pred. No.: 0.665 Length: 427  
 Score: 106.50 Matches: 79  
 Percent Similarity: 33.33% Conservative: 32  
 Best Local Similarity: 23.72% Mismatches: 90

Mon Sep 13 10:28:13 2004

Search completed: September 10, 2004, 15:41:24  
Job time : 87.5 secs

```

Query Match:      2.20%      Indels:      133
DB:              1          Gaps:        18

US-10-039-272-1 (1-2661) x IRF3_HUMAN (1-427)

QY 383 AGATGCCAATGGCAACATAGTCTATGAGAACTGCAGAAATGAG---GCTGGTTTATC 439
DB 35 ArgileProTrpIysHisGlyLeuArgGlnAspAlaGlnGln-GluaspheGlyile-- 53
QY 440 TGTGTATCCATATGTTTACAACTGGACAGCAGTGTGTGAGGACAGTGCAGGGGAAAATGG 499
DB 54 -----PheGlnAlaTrpAlaGlu-----Al 60
QY 500 CACCGGCCAAAGCCATCAATACGCTTCCCTGATGGAAACCTTTTCCTCACCACCCCGG 559
DB 60 aThrGly-----AlaTyrValProGlyArgAspLysProAspLeuProTh 75
QY 560 ATGGAGAGATGGNAATTCAT-----CTACGTCCTCCACA 594
DB 75 rTrpLysArg---AsnPheArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaGl 94
QY 595 CACTTGCTCAGTATTCCCA-----GAAATTGGGACGATGTTTCAGTGAGAGTTT 642
DB 94 uAspArgSerLysAspProHisAspProHisLysIleTyrGluPheValAsnSerGlyVa 114
QY 643 CTGTGAACACACCAATGTGACACTTGGGCTCACTCATGGAAGTGACTCTCTACAGAA 702
DB 114 lGlyAspPheSerGlnProAspThrSerProAspThrAsnGlyGly----- 129
QY 703 GACATGGNCGGCATATGTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAAACAGATC 762
DB 130 -----GlySerThrSe 133
QY 763 AGATTTCCTGTGTTGTGACTATGTTCCAGAAACGATCGAAATTCATCCGACGAAACCT 822
DB 133 rAspThrGlnGlu---AspIleLeuAspGluLeuLeuGlyAsnMetValLeuAlaProLe 152
QY 823 TCCCAAAGATCTCCCATTTATGTTGATGTCCTGATTCATGATCCTAGCCACTTCTCAA 882
DB 152 uPro-----AspPro----- 155
QY 883 TTATTCTACCATTAACATACAGTGGAGCTTCGGGGATATACTGGCCTGTTTGTTCAC 942
DB 155 ----- 155
QY 943 CAATCATACTGTGAATCACACGTATGTCTCAATGGAACCTTCAGCCTTAACTCAGTGT 1002
DB 156 -----GlyProSerLeu----- 160
QY 1003 GAAAGTCGACGACGAGGACCTTGTCCGCCACGCGCCACCACCCAGACCTTCAAACCC 1062
DB 161 -----AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnPr 179
QY 1063 CACC-----CCTTCTTTAGGACCTGCTGTGACAAACCCCTGGAGCTGAGTAGGATCC 1116
DB 179 oThrProPheProAsnLeuGlyProSer---GluAsnProLeuLysArgLeuLeuValPr 198
QY 1117 TGATGAAAACTGCCAGATTAAACAGATATG3CCAC-----TTTCAAGC 1158
DB 198 oGlyGluGluTrpGluPheGluValThrAlaPheTyrArgGlyArgGlnValPheGlnGl 218
QY 1159 CACCATCACAATGTAGAGGAATCTTAGAGGTTACATCATCCAGATGACAGACGTCCT 1218
DB 218 nThrIleSerCysProGluGlyLeuArgLeuValGly---SerGluValGlyAspArgTh 237
QY 1219 GATGCCG-----GTGCCATGGCTGAAAGCTCCCTATATAGAC----- 1255
DB 237 rLeuProGlyTrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyVa 257
QY 1256 -----TTTGTGTCGACCTGCCAAGG 1276
DB 257 lMetSerTyrValArgHisValLeuSerCysLeuGly 269

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 9, 2004, 20:27:32 ; Search time 16423 Seconds  
(without alignments)  
7022.822 Million cell updates/sec

Title: US-10-039-272-1

Perfect score: 2661

Sequence: 1 cggcagagggccagagga.....aaaaaaaaaaaaaaaaaaaaa 2661

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

15: em.ba.\*

16: em.fun.\*

17: em.hum.\*

18: em.in.\*

19: em.mu.\*

20: em.om.\*

21: em.or.\*

22: em.ov.\*

23: em.pat.\*

24: em.ph.\*

25: em.pl.\*

26: em.ro.\*

27: em.sts.\*

28: em.un.\*

29: em.vi.\*

30: em.htg.hum.\*

31: em.htg.inv.\*

32: em.htg.other.\*

33: em.htg.mus.\*

34: em.htg.pln.\*

35: em.htg.rod.\*

36: em.htg.mam.\*

37: em.htg.vrt.\*

38: em.sy.\*

39: em.htgo.hum.\*

40: em.htgo.mus.\*

41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	2650	99.6	2662	9	AF322909 Homo sapi
2	2547.8	95.7	2669	6	AR018808 Sequence
3	2547.8	95.7	2669	6	AR018814 Sequence
4	2547.8	95.7	2669	6	AX818137 Sequence
5	2547.8	95.7	2669	9	HSNMB
6	2522	94.8	2787	9	BC032783
7	2485	93.4	2683	6	AX358788 Sequence
8	2485	93.4	2683	6	AX362281 Sequence
9	2485	93.4	2683	9	AY359124 Homo sapi
10	1671	62.8	1683	6	AX677738 Sequence
11	1051.2	39.5	2282	10	BC061725
12	1051.2	39.5	2303	6	BD062749 Modulator
13	1049.8	39.5	2279	10	AF322054 Mus muscu
14	1049.8	39.5	2299	10	BC026375
15	1047.8	39.4	2213	6	AR156839 Sequence
16	1047.8	39.4	2213	6	BD269857
17	1046.4	39.3	2320	10	AF184983 Rattus no
18	1045	39.3	2305	10	MMU251685
19	1002.6	37.7	169739	9	AC005082 Homo sapi
20	988.2	37.1	221255	9	AC145883
21	854.2	32.1	898	11	G26743 human STS S
22	584.6	22.0	2467	5	CJQNR71
23	572	21.5	1690	9	BC011595 Homo sapi
24	568.8	21.4	1593	9	HSAS05015
25	546.4	20.5	621	9	BT007074 Homo sapi
26	546.4	20.5	621	12	BT007499
27	545.6	20.5	619	6	AR176437
28	545.6	20.5	619	6	BD275711
29	545.6	20.5	619	6	BD275724
30	545.6	20.5	619	6	AR220496 Sequence
31	545.6	20.5	619	6	AR220509 Sequence
32	545.6	20.5	619	6	AR255490 Sequence
33	545.6	20.5	619	6	AR255503 Sequence
34	545.6	20.5	619	6	AR281060 Sequence
35	545.6	20.5	619	6	AR281073 Sequence
36	545.6	20.5	619	6	AR437851 Sequence
37	545.6	20.5	619	6	AR437864 Sequence
38	545.6	20.5	619	6	AX365712 Sequence
39	545.6	20.5	619	6	AX365728 Sequence
40	545.6	20.5	619	6	BD226040 Compound
41	545.6	20.5	619	6	BD226053 Compound
42	537.2	20.2	572	6	AX360616 Sequence
43	526.2	19.8	570	6	AX360615 Sequence
44	518	19.5	158105	2	AC023375 Homo sapi
45	514.8	19.3	537	6	AX360618 Sequence

# ALIGNMENTS

RESULT 1  
AF322909  
LOCUS Homo sapiens transmembrane glycoprotein HGFN mRNA, complete cds.  
DEFINITION AF322909 2662 bp mRNA linear PRI 23-APR-2003  
ACCESSION AF322909  
VERSION AF322909.1 GI:11993664  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2662)  
AUTHORS Bandari, P.S., Qian, J., Yehia, G., Joshi, D.D., Maloof, P.B.,  
Potluri, J., Oh, H.S., Gascon, P., Harrison, J.S. and Rameshwar, P.  
TITLE Hematopoietic growth factor inducible neurokinin-1 type: a

transmembrane protein that is similar to neurokinin 1 interacts with substance P  
Regul. Pept. 111 (1-3), 169-178 (2003)  
22498106  
12609785  
2 (bases 1 to 2662)  
Rameshwar, P.  
Direct Submission  
Submitted (20-NOV-2000) Medicine, UMDNJ-New Jersey Medical School,  
185 South Orange Ave. MSB. Rm. E-579, Newark, NJ 07103, USA

```

Location/Qualifiers
1. .2662
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="bone marrow peripheral blood"
60. 1742
/notes="hematopoietic growth factor-inducible neurokinin-1

```

```

protein"
/codon_start=1
/product="transmembrane glycoprotein HGFIN"
/protein_id="AAG42839.1"
/db_xref="GI:11993665"
/translation="MECLYVFLGLLLAARLPDAAKRPHDVLGNRPSAYMEHNLQ
NGWSSDENNEKLYPWWRKDRNRKSNKGGVRQAVLSDSPALVGSNTTFVNNILP
PRCKEDANGNIYYENKREAGLSADPYVYNTAMSEDSGDEGTQSGSHVFPDGK
PPHPPGKRWNRNIYYFTLIGQYFKLGRCSVSVSYNTAVTILGPQLVYVRRHGR
AVPYIAQKDWVAVTQDIPVFTMGKNDNRNSDETFKDLPIMFVLIHDPSHFYLN
STINYKSGDNTGLSVSNTHYVNTHTYVLTNGTFSNLTVKAAAGPCPPPPPPPRSK
PTPLSGDPVWUELSRIPDENECQINRYGHFQATITIVEGILFVNLIQMTDVLMPV
PSSLLDFVVTGCGSIEPTVECTIIISDPTCEITQVCSVDVDMCLLTIVRTFNGSG
TYCWNLTGDDTSLATLSITVSPDRPASPFLRMANSLISVGLCAIFVTVISLLVYK
KKEVNPINENSGNVRSGLSVFLNRAKAVFFPGNOEKDPLLNKQDFGKGS"

```

ch	99.6%;	Score 2650;	DB 9;	Length 2662;
1	Similarity 100.0%;	Pred. No. 0;		
661;	Conservative 0;	Mismatches 0;	Indels 1;	Gaps 1;
1	CGGCACGAGGCCGAGAGGAATAGTTAACTTTGGTGCCTCGGTCCTGAGAATTCAGCA	60		
1	CGGCACGAGGCCGAGAGGAATAGTTAACTTTGGTGCCTCGGTCCTGAGAATTCAGCA	60		
61	TGGAATGTCCTACTATTTCCTCGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGTATG	120		
61	TGGAATGTCCTACTATTTCCTCGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGTATG	120		
121	CCGCCAAACGATTTTCATGATGTCTGGCGCAATGAAAGACCTTCTGCTTACATGAGGGAGC	180		
21	CCGCCAAACGATTTTCATGATGTCTGGCGCAATGAAAGACCTTCTGCTTACATGAGGGAGC	180		
81	ACAATCAATTAAATGGCTGGTCTTCTGATGAAATGACTGGAATGAAAACTCTACCCAG	240		
81	ACAATCAATTAAATGGCTGGTCTTCTGATGAAATGACTGGAATGAAAACTCTACCCAG	240		
441	TGTGGAAACGGGGAGACATGAGGTGGAAAAAACTCCTGGAAAGGAGGCCGTGTGACGGCG	300		
441	TGTGGAAACGGGGAGACATGAGGTGGAAAAAACTCCTGGAAAGGAGGCCGTGTGACGGCG	300		
001	TCCTGACACGTGACTCACCCAGCCCTCGTGGGCTCAAAATATAACATTTCCGGGTGAACCTGA	360		
001	TCCTGACACGTGACTCACCCAGCCCTCGTGGGCTCAAAATATAACATTTCCGGGTGAACCTGA	360		
61	TATTCCCTAGATGCCAAAGGAAGATGCAATGGCAACATAGTCTCTATGAGAAGAACTGCA	420		
61	TATTCCCTAGATGCCAAAGGAAGATGCAATGGCAACATAGTCTCTATGAGAAGAACTGCA	420		
121	GAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAGG	480		
121	GAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAGG	480		
481	ACAGTGACGGGAAATGGCACCGGCCAAAGCCATCAACGCTCTTCCTGATGGGAAC	540		

```
QY 1620 GGAATCTGGTCAAGCAAGGCGCTGAGTGTCTTTCTCAACCGTGCAGGCGGTCTTCT 1679
Db 1621 GGAATCTGGTCAAGCAAGGCGCTGAGTGTCTTTCTCAACCGTGCAGGCGGTCTTCT 1680
QY 1680 TCCCGGAAACACAGGAAAGGATCCGCTACTCAAAACCAAGAAATTTAAAGGAGTTCTT 1739
Db 1681 TCCCGGAAACACAGGAAAGGATCCGCTACTCAAAACCAAGAAATTTAAAGGAGTTCTT 1740
QY 1740 AAATTCGACCTTGTTCGAGGCTCACATTTTCAGTGCATTCAGTGCAGATGCTGGA 1799
Db 1741 AAATTCGACCTTGTTCGAGGCTCACATTTTCAGTGCATTCAGTGCAGATGCTGGA 1800
QY 1800 GTGGCTATTAACTTTTTTCTTAAAGATTAATTTAAATAGATATGTTGGTTGGGAA 1859
Db 1801 GTGGCTATTAACTTTTTTCTTAAAGATTAATTTAAATAGATATGTTGGTTGGGAA 1860
QY 1860 GTTGAATTTTATAGGTTAAATGCTAATTTAGATGGGAGAGGATTAATCTGCAGG 1919
Db 1861 GTTGAATTTTATAGGTTAAATGCTAATTTAGATGGGAGAGGATTAATCTGCAGG 1920
QY 1920 CAGCTTCAGCCATGTTGTGAAGCTCATAAAGCAACTTAGCAAGCTTTCTTTTCAATTT 1979
Db 1921 CAGCTTCAGCCATGTTGTGAAGCTCATAAAGCAACTTAGCAAGCTTTCTTTTCAATTT 1980
QY 1980 TTTATGTTTCACTTATAAAGTCTTAGGTAAGTACTAGATAGAAACACTGTGCCGAG 2039
Db 1981 TTTATGTTTCACTTATAAAGTCTTAGGTAAGTACTAGATAGAAACACTGTGCCGAG 2040
QY 2040 AGTAAGGAGAAAGCTACTATTGATTAAGGCTTAACCCAGGTTAACTGCAAGAGAGGCG 2099
Db 2041 AGTAAGGAGAAAGCTACTATTGATTAAGGCTTAACCCAGGTTAACTGCAAGAGAGGCG 2100
QY 2100 GGATACCTTCAGCTTTCATGTAAGTCTGATGATGATGATGATGATGATGATGATGAT 2159
Db 2101 GGATACCTTCAGCTTTCATGTAAGTCTGATGATGATGATGATGATGATGATGATGAT 2160
QY 2160 ATCATGTTTCAAGCTAATCCCACTTCAATACACACTCATGAACTCTGATGAAAC 2219
Db 2161 ATCATGTTTCAAGCTAATCCCACTTCAATACACACTCATGAACTCTGATGAAAC 2220
QY 2220 AATAAGGCGCAAGCCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2279
Db 2221 AATAAGGCGCAAGCCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280
QY 2280 CTCTCAATATGGTGGGAGTATTTGGTGACACCTACTTTGCTGGCTGAGTGAAGGA 2339
Db 2281 CTCTCAATATGGTGGGAGTATTTGGTGACACCTACTTTGCTGGCTGAGTGAAGGA 2340
QY 2340 ATGATATTCATATTCATTTATTCATGGACATTTAGTTAGTGTCTTTTATATACAGG 2399
Db 2341 ATGATATTCATATTCATTTATTCATGGACATTTAGTTAGTGTCTTTTATATACAGG 2400
QY 2400 CATGATGCTGAGTGACACTCTTGATATTTCCAAATTTTGTATGCTGTCACATA 2459
Db 2401 CATGATGCTGAGTGACACTCTTGATATTTCCAAATTTTGTATGCTGTCACATA 2460
QY 2460 TTTGAATCAAAATATTAAGACTTTCCAAAATTTGGTCCCTGGTTTTTATGGCAACTT 2519
Db 2461 TTTGAATCAAAATATTAAGACTTTCCAAAATTTGGTCCCTGGTTTTTATGGCAACTT 2520
QY 2520 GATCAGTAAGGATTTCCCTCTGTTTGGAACTTAAACCACTTACTATATGTAGACAAG 2579
Db 2521 GATCAGTAAGGATTTCCCTCTGTTTGGAACTTAAACCACTTACTATATGTAGACAAG 2580
QY 2580 CATTTTTTTTTTCTTCTGAAAATTAAGGAGGAGACAAAAAATTAATTAATTAATTAATTA 2639
Db 2581 CATTTTTTTTTTCTTCTGAAAATTAAGGAGGAGACAAAAAATTAATTAATTAATTAATTA 2640
QY 2640 AAAAAAAAAAAAAAAAAAAAAA 2661
Db 2641 AAAAAAAAAAAAAAAAAAAAAA 2662
```

```
RESULT 2
AR018808
LOCUS AR018808 2669 bp DNA linear PAT 05-DEC-1998
DEFINITION Sequence 91 from patent US 5783182.
ACCESSION AR018808
VERSION AR018808.1 GI:3973922
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 2669)
TITLE Thompson, T.C.
JOURNAL Method for identifying metastatic sequences
FEATURES Patent: US 5783182-A 91 21-JUL-1998;
Location/Qualifiers
source 1..2669
/organism="unknown"
/mol_type="unassigned DNA"
```

## ORIGIN

```
Query Match 95.7%; Score 2547.8; DB 6; Length 2669;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2591; Conservative 0; Mismatches 17; Indels 3; Gaps 3;
QY 28 AACCTTGGTGGCTCGCTCCGTGAGAAATTCAGCATGGATGCTCTACTATTTCCTGGGAT 87
Db 60 ACCTTGATGCTCGCTCCGTGAGAAATTCAGCATGGATGCTCTACTATTTCCTGGGAT 119
QY 88 TTCTGTCTTGGTGGCTCGCTCCGTGAGAAATTCAGCATGGATGCTCTACTATTTCCTGGGAT 147
Db 120 TTCTGTCTTGGTGGCTCGCTCCGTGAGAAATTCAGCATGGATGCTCTACTATTTCCTGGGAT 179
QY 148 GCAATGAAGACCTTCTGCTTACATGAGGAGACAAATCAATTAATTAATTAATTAATTAAT 207
Db 180 GCAATGAAGACCTTCTGCTTACATGAGGAGACAAATCAATTAATTAATTAATTAATTAAT 239
QY 208 ATGAAATATGACTGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 267
Db 240 ATGAAATATGACTGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 299
QY 268 AAAATCTCTGGAAGGAGGCGCTGTGAGGCGGTCTCTGAGCAGTACTCAGCAGCCTCG 327
Db 300 AAAATCTCTGGAAGGAGGCGGTGTGAGGCGGTCTCTGAGCAGTACTCAGCAGCCTCG 359
QY 328 TGGGCTCAATATATAAATTTGGCGTGAACCTGATATTCCTCATGTCGCAAAAGGAAAGATG 387
Db 360 TGGGCTCAATATATAAATTTGGCGTGAACCTGATATTCCTCATGTCGCAAAAGGAAAGATG 419
QY 388 CCAATGGCAACATAGTCTATGAGAAAGAACTGAGAAATGAGGCTGCTTTATCTGCTGATC 447
Db 420 CCAATGGCAACATAGTCTATGAGAAAGAACTGAGAAATGAGGCTGCTTTATCTGCTGATC 479
QY 448 CATATGTTTAACTGGACAGCATGGTCAAGAGACAGTACGCGGGAATGACCGGCG 507
Db 480 CATATGTTTAACTGGACAGCATGGTCAAGAGACAGTACGCGGGAATGACCGGCG 539
QY 508 AAAGCCATCATAACTCTTCCCTGATGGGAAACCTTTTCTCACCACCCCGGATGAGAA 567
Db 540 AAAGCCATCATAACTCTTCCCTGATGGGAAACCTTTTCTCACCACCCCGGATGAGAA 599
QY 568 GATGAAATTTCACTACAGCTTTCCACACACTGGTCACTATTTCCAGAAATTTGGGACGAT 627
Db 600 GATGAAATTTCACTACAGCTTTCCACACACTGGTCACTATTTCCAGAAATTTGGGACGAT 659
QY 628 GTTCACTGAGAGTTTCTGTGAACACAGCCCAATGTGACACTTTGGGCTCAACTCATGGAG 687
Db 660 GTTCACTGAGAGTTTCTGTGAACACAGCCCAATGTGACACTTTGGGCTCAACTCATGGAG 719
QY 688 TGACTGTCTACAGAGACATGACCGGGATATGTTCCCATCGCACAAAGTGAAGATGCT 747
Db 720 TGACTGTCTACAGAGACATGACCGGGATATGTTCCCATCGCACAAAGTGAAGATGCT 779
```

Db	1860	ATTATTGTTAAATAGATAATTGTTGGTTTGGGGAAGTTGAAATTTTATATAGGTTAAATGTC	1919
Qy	1887	TTTTAGAGATGGGAGAGGAGATTATACTGACGACAGCTTCAGCCATGTTGTGAAACTGAT	1946
Db	1920	TTTTAGAGATGGGAGAGGAGATTATACTGACGACAGCTTCAGCCATGTTGTGAAACTGAT	1979
Qy	1947	AAAAGCAACTTAGCAAGGCTTCCTTTTCATTATTTTTTATGTTTTCATTATAAAGCTCTTAG	2006
Db	1980	AAAAGCAACTTAGCAAGGCTTCCTTTTCATTATTTTTTATGTTTTCATTATAAAGCTCTTAG	2039
Qy	2007	GTAACCTAGTAGAGATAGAAAACACTGTGTCCGAGAGTAAGGAGAGAAGCTACTATTGATTA	2066
Db	2040	GTAACCTAGTAGAGATAGAAAACACTGTGTCCGAGAGTAAGGAGAGAAGCTACTATTGATTA	2099
Qy	2067	GAGCCTAACCCAGGTTTAACCTGCAAGAGAGCGGGATACCTTCAGCTTTCCATGTAACTG	2126
Db	2100	GAGCCTAACCCAGGTTTAACCTGCAAGAGAGCGGGATACCTTCAGCTTTCCATGTAACTG	2159
Qy	2127	TATGCATAAAGCCCAATGTAGTCAGGTTTCTTAAGATCATGTTTCCAAGCTAACTGAATCCCA	2186
Db	2160	TATGCATAAAGCCCAATGTAGTCAGGTTTCTTAAGATCATGTTTCCAAGCTAACTGAATCCCA	2219
Qy	2187	CTTCAATACACATCATGAACTCCTGTATGGAACAAATAACAGGCCCAAGCCCTGTGGTATGA	2246
Db	2220	CTTCAATACACATCATGAACTCCTGTATGGAACAAATAACAGGCCCAAGCCCTGTGGTATGA	2279
Qy	2247	TGTGCACACTTGTCTAGACTCAGAAAAAATACTACTTCTATAAATGGGTGGGAGTATTTTG	2306
Db	2280	TGTGCACACTTGTCTAGACTCAGAAAAAATACTACTTCTATAAATGGGTGGGAGTATTTTG	2339
Qy	2307	GTGCAAACTACTTTGCTTGGCTGAGTGAAGGAATGATTAATCATATATTCATTTATTCOA	2366
Db	2340	GTGCAAACTACTTTGCTTGGCTGAGTGAAGGAATGATTAATCATATATTCATTTATTCOA	2399
Qy	2367	TGGACATTTAGTTAGTGCCTTTTATATACAGGCAATGATGCTGAGTGACACTCTTGTTGA	2426
Db	2400	TGGACATTTAGTTAGTGCCTTTTATATACAGGCAATGATGCTGAGTGACACTCTTGTTGA	2459
Qy	2427	TATTTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATCAAAATATTAAGACTTTCC	2486
Db	2460	TATTTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATCAAAATATTAAGACTTTCC	2518
Qy	2487	AAAAATTTGGTCCCTGGTTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTTTG	2546
Db	2519	AAAAATTTGGTCCCTGGTTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTTTG	2578
Qy	2547	GAACTAAACCACTTACTATATGTTAGACAAGACATTTTTTTTTTCTCTCTGTAATAA	2606
Db	2579	TACTAAACCACTTACTATATGTTAGACAAGACATTTTTTTTTTCTCTCTGTAATAA	2638
Qy	2607	AAAAATGAGGGAAGACAAAAA	2636
Db	2639	TAAAGTGTGGGAAGACAAAAA	2669
RESULT 3			
LOCUS	AR018814	2669 bp	DNA
DEFINITION	Sequence 101 from patent US 5783182.		linear
ACCESSION	AR018814		
VERSION	AR018814.1	GI:3973928	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 2669)		
AUTHORS	Thompson,T.C		
TITLE	Method for identifying metastatic sequences		
JOURNAL	Patent: US 5783182-A 101 21-JUL-1998;		
FEATURES	Location/Qualifiers		
source	1..2669		
	/organism="unknown"		

```

RESULT 3
AR018814 LOCUS 2669 bp DNA
DEFINITION Sequence 101 from patent US 5783182.
ACCESSION AR018814
VERSION AR018814.1 GI:3973928
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 2669)
TITLE Thompson,T.C
JOURNAL Method for identifying metastatic sequences
FEATURES Patent: US 5783182-A 101 21-JUL-1998;
            Location/Qualifiers
            1..2669
            /organism="unknown"
            source

```

748	QY	ACGTGTAACAGATCAGATTCCTGTGTTTCTGTGACTATGTTCCAGAGAAACGATCGAAATT	807
	Db		
780	QY	ACGTGTAACAGATCAGATTCCTGTGTTTCTGTGACTATGTTCCAGAGAAACGATCGAAATT	839
	Db		
808	QY	CATCCGACGAAACCTTCC-CAAGAGATCCCCATTATGTTTGTATGTCCTGATTCATGATC	866
	Db		
840	QY	CATCCGACGAAACCTTCCCAAGATCTCCCATTAATGTTTGTATGTCCTGATTCATGATC	899
	Db		
867	QY	CTAGCCACTTCCTCAATATTCTACCAATTAACTACAGTGGAGCTTCGGGATATAACTG	926
	Db		
900	QY	CTAGCCACTTCCTCAATATTCTACCAATTAACTACAGTGGAGCTTCGGGATATAACTG	959
	Db		
927	QY	GCCTGTTGTTTCCACCAATCATACTGTGAATCACACGTATGTGCTCAATGGAACCTTCA	986
	Db		
960	QY	GCCTGTTGTTTCCACCAATCATACTGTGAATCACACGTATGTGCTCAATGGAACCTTCA	1019
	Db		
987	QY	GCCTTAACCTCACTGTGAAAGCTGCAGCACGAGACCTTGTCCGCCACCGCCACCAACAC	1046
	Db		
1020	QY	GCCTTAACCTCACTGTGAAAGCTGCAGCACGAGACCTTGTCCGCCACCGCCACCAACAC	1079
	Db		
1047	QY	CCAGACCTTCAAAACCCACCTTCTTTAGNACCTGCTGGTGACAAACCCCTGGAGCTGA	1106
	Db		
1080	QY	CCAGACCTTCAAAACCCACCTTCTTTAGNACCTGCTGGTGACAAACCCCTGGAGCTGA	1139
	Db		
1107	QY	GTAGGATTCCTGATGAAACCTGCCAGATTACAGATATGGCCACTTTCAAGCCACCATCA	1166
	Db		
1140	QY	GTAGGATTCCTGATGAAACCTGCCAGATTACAGATATGGCCACTTTCAAGCCACCATCA	1199
	Db		
1167	QY	CAATTGTAGAGGGAATCTTTAGAGGTTAAACATCATCCAGATGACAGACGTCTGTATGCCGG	1226
	Db		
1200	QY	CAATTGTAGAGGGAATCTTTAGAGGTTAAACATCATCCAGATGACAGACGTCTGTATGCCGG	1259
	Db		
1227	QY	TGCCATGGCCTGAAAGCTCCCTAAATAGACTTTGTGTGACCTGCCAAGGAGCATTCCTCA	1286
	Db		
1260	QY	TGCCATGGCCTGAAAGCTCCCTAAATAGACTTTGTGTGACCTGCCAAGGAGCATTCCTCA	1319
	Db		
1287	QY	CGGAGGTCTGTACATCATTTCTGACCCACCTGCCAGATCACCCAGAGACACAGTCTGCA	1346
	Db		
1320	QY	CGGAGGTCTGTACCATCATTTCTGACCCACCTGCCAGATCACCCAGAGACACAGTCTGCA	1379
	Db		
1347	QY	GCCCTGTGGATGTGATGAGATGTGCTGTGATCTGTGAGAGAACTTCAATGGGTCTG	1406
	Db		
1380	QY	GCCCTGTGGATGTGATGAGATGTGCTGTGATCTGTGAGAGAACTTCAATGGGTCTG	1439
	Db		
1407	QY	GGACGTACTGTGTGAACCTCAACCTGGGGATGACCAAGCTGGCTCTCAGGAGCACCC	1466
	Db		
1440	QY	GGACGTACTGTGTGAACCTCAACCTGGGGATGACCAAGCTGGCTCTCAGGAGCACCC	1499
	Db		
1467	QY	TGAATTTCTGTTCTGTGACAGAGACCCAGCTCGCTTTAAGGATGGCAAAACAGTCCCTGA	1526
	Db		
1500	QY	TGAATTTCTGTTCTGTGACAGAGACCCAGCTCGCTTTAAGGATGGCAAAACAGTCCCTGA	1559
	Db		
1527	QY	TCTCCGTTGGCTGTTGGCCATATTGTCTACTGTGATCTCCCTCTGGTGTACAAAAAC	1586
	Db		
1560	QY	TCTCCGTTGGCTGTTGGCCATATTGTCTACTGTGATCTCCCTCTGGTGTACAAAAAC	1619
	Db		
1587	QY	ACAAGGAATACAAACCAATAGATAATGTCCTGGGAATGTGCTCAGAGCAAGGCTGGA	1646
	Db		
1620	QY	ACAAGGAATACAAACCAATAGATAATGTCCTGGGAATGTGCTCAGAGCAAGGCTGGA	1679
	Db		
1647	QY	GTGTCCTTCTCAACCGTGCAGAGCCGCTGTTCTTCCCGGAAAACAGGAAAAGGATCCGC	1706
	Db		
1680	QY	GTGTCCTTCTCAACCGTGCAGAGCCGCTGTTCTTCCCGGAAAACAGGAAAAGGATCCGC	1739
	Db		
1707	QY	TACTCAAAAAACCAAGAAATTTAAAGGAGTTTCTTAAATTCGACCTTGTTCGAAGCTCA	1766
	Db		
1740	QY	TACTCAAAAAACCAAGAAATTTAAAGGAGTTTCTTAAATTCGACCTTGTTCGAAGCTCA	1799
	Db		
1767	QY	CTTTTCAGTGCCATTGATGTGAGATGTGCTGGAGTGGCTATTAACTTTTTCCTTAAG	1826
	Db		
1800	QY	CTTTTCAGTGCCATTGATGTGAGATGTGCTGGAGTGGCTATTAACTTTTTCCTTAAG	1859
	Db		



```

/mol_type="unassigned DNA"

```

Query Match 95.7%; Score 2547.8; DB 6; Length 2669;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 2591: Conservative

	Matches 2591;	Conservative	0;	Mismatches	17;	Indels	3;	Gaps	3;
QY	28	AACTTGGTGGCTGGTCCGTGAGAAATTCAGATGGAATGTCTCTACATATTTCTTGGGAT	87						
Db	60	ACCTTGAGTGGCTGGTCCGTGAGAAATTCAGATGGAATGTCTCTACATATTTCTTGGGAT	119						
QY	88	TTCTGCTCTGGCTGCAAGATTGCCACTTGATGCCGCCAAACGATTTCAATGATGTGCTGG	147						
Db	120	TTCTGCTCTGGCTGCAAGATTGCCACTTGATGCCGCCAAACGATTTCAATGATGTGCTGG	179						
QY	148	GCAATGAAGACCTTCTGCTTACATGAGGAGCACAATCAATTAATTAATGGCTGTGCTTCTG	207						
Db	180	GCAATGAAGACCTTCTGCTTACATGAGGAGCACAATCAATTAATTAATGGCTGTGCTTCTG	239						
QY	208	ATGAAATGACTTGAATGAAAACTCTACCCAGTGTGGAACGGGGAGACATGAGGTGGA	267						
Db	240	ATGAAATGACTTGAATGAAAACTCTACCCAGTGTGGAACGGGGAGACATGAGGTGGA	299						
QY	268	AAAACTCTCGAAGGGAGCGGTGTGAGGCGGTCTGACCAATGACTCACCAAGCCCTCG	327						
Db	300	AAAACTCTCGAAGGGAGCGGTGTGAGGCGGTCTGACCAATGACTCACCAAGCCCTCG	359						
QY	328	TGGGCTCAAAATATAACATTTCCGGTGAACCTGTATTTCCCTAGATGCCAAAAGAAATG	387						
Db	360	TGGGCTCAAAATATAACATTTCCGGTGAACCTGTATTTCCCTAGATGCCAAAAGAAATG	419						
QY	388	CCAAATGGCAACATAGTCTATCAGAAAGAACTGCACAAATGAGGCTGTTTATCTCTGATC	447						
Db	420	CCAAATGGCAACATAGTCTATCAGAAAGAACTGCACAAATGAGGCTGTTTATCTCTGATC	479						
QY	448	CATATGTTTACAATGGACACATGTGTGAGAGGACGTGACGGGGGAAAAATGGCACCGGCC	507						
Db	480	CATATGTTTACAATGGACACATGTGTGAGAGGACGTGACGGGGGAAAAATGGCACCGGCC	539						
QY	508	AAAGCCATCATAGCTCTTCCGTATGGGAAAACCTTTTCTCCACCCCGGATGAGAA	567						
Db	540	AAAGCCATCATAGCTCTTCCGTATGGGAAAACCTTTTCTCCACCCCGGATGAGAA	599						
QY	568	GATGGAATTTCACTACGTCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGAT	627						
Db	600	GATGGAATTTCACTACGTCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGAT	659						
QY	628	GTTCACTGAGATTTCTGTGAAACACGCCAAATGTGACATTTGGGCTTCAACTCATGGAAG	687						
Db	660	GTTCACTGAGATTTCTGTGAAACACGCCAAATGTGACATTTGGGCTTCAACTCATGGAAG	719						
QY	688	TGACTGTCTACAGACACATGGAGGGGCATATGTTCCCATGCGCAAAAGTGTGT	747						
Db	720	TGACTGTCTACAGACACATGGAGGGGCATATGTTCCCATGCGCAAAAGTGTGT	779						
QY	748	ACGTGTTTACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAGACGATTCGAATT	807						
Db	780	ACGTGTTTACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAGACGATTCGAATT	839						
QY	808	CATCCGACGAACTTCC--AAAGATCTCCCCATTAATGTTGATGTCTGATTCATGATC	866						
Db	840	CATCCGACGAACTTCTCCAAAGATCTCCCCATTAATGTTGATGTCTGATTCATGATC	899						
QY	867	CTAGCCACTTCTTCAATTAATCTTACCATTAACTACAGTGGAGCTTCGGGGATAATCTG	926						
Db	900	CTAGCCACTTCTTCAATTAATCTTACCATTAACTACAGTGGAGCTTCGGGGATAATCTG	959						
QY	927	GCTGTTTGTTCACCAATCATCTGTGAATCACACGATATCTGTCTCAATGGAACCTTCA	986						
Db	960	GCTGTTTGTTCACCAATCATCTGTGAATCACACGATATCTGTCTCAATGGAACCTTCA	1019						
QY	987	GCCTTAACTCTACTGTGAAGCTGCAGACACGAGCCTTGTCCGCCACCGCCACACACAC	1046						

QY	2127	TATGATATAAGCCAAATAGTCCAGTTCTTAAGTCAATGTTCCAGCTAATCAATCCCA	2186
Db	2160	TATGATATAAGCCAAATAGTCCAGTTCTTAAGTCAATGTTCCAGCTAATCAATCCCA	2219
QY	2187	CTTCAATACACACTCATGAATCTCTGATGGAACAATAACAGGCCCAAGCCCTGTGGTATGA	2246
Db	2220	CTTCAATACACACTCATGAATCTCTGATGGAACAATAACAGGCCCAAGCCCTGTGGTATGA	2279
QY	2247	TGTGCACACTTGTCTAGACTCAGAAAAAATACTCTCATTAATGGGTGGAGTATTG	2306
Db	2280	TGTGCACACTTGTCTAGACTCAGAAAAAATACTCTCATTAATGGGTGGAGTATTG	2339
QY	2307	GTGACACACTTGTCTAGACTCAGAAAAAATACTCTCATTAATGGGTGGAGTATTG	2366
Db	2340	GTGACACACTTGTCTAGACTCAGAAAAAATACTCTCATTAATGGGTGGAGTATTG	2399
QY	2367	TGGACATTTAGTGTAGTCTTTTATATACAGGAGATGCTGAGTGACACTCTTGTGTA	2426
Db	2400	TGGACATTTAGTGTAGTCTTTTATATACAGGAGATGCTGAGTGACACTCTTGTGTA	2459
QY	2427	TATTTCCAAATTTGTATAGTCGCTGCACATATTTGAAATCAAAATTAAGACTTTCC	2486
Db	2460	TATTTCCAAATTTGTATAGTCGCTGCACATATTTGAAATC-ATATATTAAGACTTTCC	2518
QY	2487	AAAAATTTGTCCTGTTTTCATGCAACTGATGATGAGTAAAGTATTTCCCTCTGTTG	2546
Db	2519	AAAGATGAGTCCCTGTTTTCATGCAACTGATGATGAGTAAAGTATTTCCCTCTGTTG	2578
QY	2547	GAACTAAACCAATTTACTATATGTTAGACAAGACATTTTTTTTTTCTCTCTGAAAAA	2606
Db	2579	TAACTAAACCAATTTACTATATGTTAGACAAGACATTTTTTTTTTCTCTCTGAAAAA	2638
QY	2607	-AAAAATGAGGAGAGACAAAAAATAAAAAA	2636
Db	2639	TAAAGTGTGGAAGAGACAAAAAATAAAAAA	2669
RESULT 4			
AX818137			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 2591; Conservative			
QY	28	AACCTTGGTGGCTGCGCTCCGTGAGAAATTCAGCATGGAATGCTCTACTATTCTCTGGAT	87
Db	60	ACCTTGGTGGCTGCGCTCCGTGAGAAATTCAGCATGGAATGCTCTACTATTCTCTGGAT	119
QY	88	TTCTGCTCTGGCTGCAAGATTTGACCTTATGCGGCAACAGATTTTCATGATGCTGG	147
Db	120	TTCTGCTCTGGCTGCAAGATTTGACCTTATGCGGCAACAGATTTTCATGATGCTGG	179

QY	148	GCAATGAAAGACCTTCTGCTTACATGAGGAGGACAAATCAATTAATGAGCTGCTCTCTG	207
Db	180	GCAATGAAAGACCTTCTGCTTACATGAGGAGGACAAATCAATTAATGAGCTGCTCTCTG	239
QY	208	ATGAAATGACTGGAATGAAATCTCTACCCAGTGTGAAAGCGGGAGACATGAGGTGGA	267
Db	240	ATGAAATGACTGGAATGAAATCTCTACCCAGTGTGAAAGCGGGAGACATGAGGTGGA	299
QY	268	AAAATCTCTGGAAGGAGGCGGTGTGAGGCGGTCTGACAGTGAATCACCAGCCCTCG	327
Db	300	AAAATCTCTGGAAGGAGGCGGTGTGAGGCGGTCTGACAGTGAATCACCAGCCCTCG	359
QY	328	TGGGCTCAAAATTAACATTTTCCGCTGAACCTGATATTTCCCTAGATGCCAAAGAGATG	387
Db	360	TGGGCTCAAAATTAACATTTTCCGCTGAACCTGATATTTCCCTAGATGCCAAAGAGATG	419
QY	388	CCATGCGCAATAGTCTTATGAGAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC	447
Db	420	CCATGCGCAATAGTCTTATGAGAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC	479
QY	448	CATATGTTTACAACTGGACAGCATGGTCAAGAGACAGTGACGGGGAATAATGGCACCGCC	507
Db	480	CATATGTTTACAACTGGACAGCATGGTCAAGAGACAGTGACGGGGAATAATGGCACCGCC	539
QY	508	AAAGCATATAACGCTCTTCCCTGATGGGAAACCTTTTCCCTCACACCCCGGATGGAGAA	567
Db	540	AAAGCATATAACGCTCTTCCCTGATGGGAAACCTTTTCCCTCACACCCCGGATGGAGAA	599
QY	568	GATGGAATTTCTATCTACCTCTTCCACACACTTTGGTCACTATTTCCAGAAATTTGGACCAT	627
Db	600	GATGGAATTTCTATCTACCTCTTCCACACACTTTGGTCACTATTTCCAGAAATTTGGACCAT	659
QY	628	GTTCACTGAGAGATTTCTGTGAACACAGCCAAATGTGACACTTTGGGCGCTCAACTCATGGAAG	687
Db	660	GTTCACTGAGAGATTTCTGTGAACACAGCCAAATGTGACACTTTGGGCGCTCAACTCATGGAAG	719
QY	688	TGACTGTCTACAGAGACATGAGCGGATATGTTCCCATCCACAAAGTGAAGATGTCT	747
Db	720	TGACTGTCTACAGAGACATGAGCGGATATGTTCCCATCCACAAAGTGAAGATGTCT	779
QY	748	ACGTGTAACACATCAGATTTCTGTGTTGTGACTATGTTCCAGAGAAACGATCGAAAT	807
Db	780	ACGTGTAACACATCAGATTTCTGTGTTGTGACTATGTTCCAGAGAAACGATCGAAAT	839
QY	808	CATCCGAGAAACCTTCC-CAAAGATCTCCCATTTATGTTGATGCTCTGATTCATGATC	866
Db	840	CATCCGAGAAACCTTCC-CAAAGATCTCCCATTTATGTTGATGCTCTGATTCATGATC	899
QY	867	CTAGCCACTTCTCAATTTCTACCATTAATCTACAAGTGGAGCTTCGGGGATAATACTG	926
Db	900	CTAGCCACTTCTCAATTTCTACCATTAATCTACAAGTGGAGCTTCGGGGATAATACTG	959
QY	927	GCCTGTTCTTCCCAATCATCTGTAATCACACATGATGCTCAATGGAACCTTCA	986
Db	960	GCCTGTTCTTCCCAATCATCTGTAATCACACATGATGCTCAATGGAACCTTCA	1019
QY	987	GCCTTAACCTCACTGTGAAAGCTGACGACACAGGACCTTGTGCGGACACCGCCACCCAC	1046
Db	1020	GCCTTAACCTCACTGTGAAAGCTGACGACACAGGACCTTGTGCGGACACCGCCACCCAC	1079
QY	1047	CGAGACTTCAAAACCCCTTCTTTAGGACCTGCTGTGACAAACCCCTGAGCTGA	1106
Db	1080	CGAGACTTCAAAACCCCTTCTTTAGGACCTGCTGTGACAAACCCCTGAGCTGA	1139
QY	1107	GTAGGATTTCTGTATGAAATCTGCCAGATTAACAGATATGCGCACTTTCAAGCCACCATCA	1166
Db	1140	GTAGGATTTCTGTATGAAATCTGCCAGATTAACAGATATGCGCACTTTCAAGCCACCATCA	1199
QY	1167	CAATTTGAGAGGAAATCTTAGAGGTTAAATCATCCAGATGACAGAGCTCTGATGCGG	1226
Db	1200	CAATTTGAGAGGAAATCTTAGAGGTTAAATCATCCAGATGACAGAGCTCTGATGCGG	1259
QY	1227	TGCCATGGCTGAAAGCTCCCTAATAGACTTTTGTGCTGACCTGTCAGAGGAGCATTCCTCA	1286

Db	2146	GAAGAGCGGGGATCTTTTCAGCTTTCCATGTAACCTGTATGTCATTAAGCCAACTGATGCCA	2205
QY	2151	GTTCCTTAAGATCATGTTTCAAGCTACTGAATCCCACTTCAATACACACTCATGAACTCC	2210
Db	2206	GTTCCTTAAGATCATGTTTCAAGCTACTGAATCCCACTTCAATACACACTCATGAACTCC	2265
QY	2211	TGATGGAACAATAACAGGCCCCAAGCCTGTGGTATGATGTGCACACTTGTAGACTCAGAA	2270
Db	2266	TGATGGAACAATAACAGGCCCCAAGCCTGTGGTATGATGTGCACACTTGTAGACTCAGAA	2325
QY	2271	AAAACTACTCTCTATAAATGGGTGGGAGTATTTGGTGACAACCTACTTTGCTTGGCTG	2330
Db	2326	AAAACTACTCTCTATAAATGGGTGGGAGTATTTGGTGACAACCTACTTTGCTTGGCTG	2385
QY	2331	AGTGAAGGAATGATATTCATATATTCATTTATTCATGACACATTTAGTTAGTCTTTTAA	2390
Db	2386	AGTGAAGGAATGATATTCATATATTCATTTATTCATGACACATTTAGTTAGTCTTTTAA	2445
QY	2391	TATACCGGCGATGATGCTGAGTGACACTCTTGTGTATATTTCCAAATTTTCTATAGTCG	2450
Db	2446	TATACCGGCGATGATGCTGAGTGACACTCTTGTGTATATTTCCAAATTTTGTACAGTCG	2505
QY	2451	CTGCACATATTTGAAATCAAAATATTAAGACTTTTCCAAAAATTTGGTCCCTGGTTTTC	2510
Db	2506	CTGCACATATTTGAAATC-ATATATTAAGACTTTTCCAAAGATGAGGTCCCTGGTTTTC	2564
QY	2511	TGGCAACTTCATCAGTAAAGGATTTCCCTCTGTTTGGAACTTAAACCACTTACTATATGT	2570
Db	2565	TGGCAACTTCATCAGTAAAGGATTTCCCTCTGTTTGGAACTTAAACCACTTACTATATGT	2624
QY	2571	TAGACAAGACATTTTTTTTTTTTCTCTCTGTAATAA-AAAAATGAGGGAAGACAAAAA	2629
Db	2625	TAGACAATGACATCTTTTCTCTCTCTCAAAAATAAGTGTGGGAAGACAAAAA	2684
QY	2630	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	2661
Db	2685	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	2716
RESULT 7			
LOCUS	AX358788		
DEFINITION	Sequence 41 from Patent WO0193983.	DNA	linear
ACCESSION	AX358788		
VERSION	AX358788.1	GI:18675295	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.		
TITLE	Baker, K.P., Desnoyers, L., Gerritsen, M.E., Goddard, A.,		
JOURNAL	Godowski, P.J., Grimaldi, J.C., Gurney, A.L., Smith, V., Stephan, J.P.,		
FEATURES	Secreted and transmembrane polypeptides and nucleic acids encoding		
source	the same		
Location/Qualifiers	Patent: WO 0193983-A 41 13-DEC-2001;		
1.2683	Genentech Inc. (US)		
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
ORIGIN			
Query Match	93.4%	Score 2485;	DB 6; Length 2683;
Best Local Similarity	97.8%	Pred. No. 0;	
Matches 2576;	Conservative	0; Mismatches 20;	Indels 39; Gaps 4;
QY	28	AACCTTGTGTGCTCGGTCCGTGAGAAATTCAGCATGGAATGCTCTACTATTTTCCTGGGAT	87
Db	50	ACCTTGTGTGCTCGGTCCGTGAGAAATTCAGCATGGAATGCTCTACTATTTTCCTGGGAT	109

QY	1131	AGATTAA	CAGATATGGCCACTTTTCAAGCCACCATCAAA	TGTAGAGGAA	TCTTAGAGG	1190
DB						
QY	1190	AGATTAA	CAGATATGGCCACTTTTCAAGCCACCATCAAA	TGTAGAGGAA	TCTTAGAGG	1249
DB						
QY	1191	TTAACA	TATCATCAGATGACAGAGCTCTGATGCCGGTGCCATGGCTGAAAGCTCCCTAA			1250
DB						
QY	1250	TTAACA	TATCATCAGATGACAGAGCTCTGATGCCGGTGCCATGGCTGAAAGCTCCCTAA			1309
DB						
QY	1251	TAGACTTT	TGTCGTGACCTCCAAAGGAGCATTCACACGGAGGTCTGTACCATCATTTCTG			1310
DB						
QY	1310	TAGACTTT	TGTCGTGACCTCCAAAGGAGCATTCACACGGAGGTCTGTACCATCATTTCTG			1369
DB						
QY	1311	ACCCAC	CTGCCAGATCACCCAGAAACACAGTCTGACAGCCCTGTGGATGTGGATGAATGT			1370
DB						
QY	1370	ACCCAC	CTGCCAGATCACCCAGAAACACAGTCTGACAGCCCTGTGGATGTGGATGAATGT			1429
DB						
QY	1371	GTCTGT	CTGACTGTGAGACGAACTTCAATGGGTCTGGHAGTACTGTGTAACCTCACCC			1430
DB						
QY	1430	GTCTGT	CTGACTGTGAGACGAACTTCAATGGGTCTGGHAGTACTGTGTAACCTCACCC			1489
DB						
QY	1431	TGGGGAT	GACACAAGCTGGCTCTCACGAGACCCCTGATTTCTGTTCCTGCACAGAGCC			1490
DB						
QY	1490	TGGGGAT	GACACAAGCTGGCTCTCACGAGACCCCTGATTTCTGTTCCTGCACAGAGCC			1549
DB						
QY	1491	CAGCCT	CGCCTTTAAGGATGGCAAAACAGTCCCTGTATCTCCGTGGCTGCTTGGCCATAT			1550
DB						
QY	1550	CAGCCT	CGCCTTTAAGGATGGCAAAACAGTCCCTGTATCTCCGTGGCTGCTTGGCCATAT			1609
DB						
QY	1551	TTGTCT	CTGTGATCTCCCTTGTGTACAAAACACACAGGATACAA	CCCAATAGAA		1610
DB						
QY	1610	TTGTCT	CTGTGATCTCCCTTGTGTACAAAACACACAGGATACAA	CCCAATAGAA		1669
DB						
QY	1611	ATAGTCT	TGGAAATGGTTCAGAAACAAAGCCCTGAGTGTCTTTCTCAACCGTGCAAAAG			1670
DB						
QY	1670	ATAGTCT	TGGAAATGGTTCAGAAACAAAGCCCTGAGTGTCTTTCTCAACCGTGCAAAAG			1729
DB						
QY	1671	CCGTGT	CTTCTCCCGGAAACAGGAAAGGATCGCTACTCAAAAACCAAGAATTTAAAG			1730
DB						
QY	1730	CCGTGT	CTTCTCCCGGAAACAGGAAAGGATCGCTACTCAAAAACCAAGAATTTAAAG			1789
DB						
QY	1731	GAGTTT	CTTAAAATTTGACACTTTTCTGAAGCTCACTTTTCAGTGCCATTGATGTGAGA			1790
DB						
QY	1790	GAGTTT	CTTAAAATTTGACACTTTTCTGAAGCTCACTTTTCAGTGCCATTGATGTGAGA			1849
DB						
QY	1791	TGTGCT	GGGTGGCTATTAACTTTTTTCTCTAAAGATTATTTGTTAAATAGATATTGTGG			1850
DB						
QY	1850	TGTGCT	GGGTGGCTATTAACTTTTTTCTCTAAAGATTATTTGTTAAATAGATATTGTGG			1909
DB						
QY	1851	TTTTGG	GAAGTTGAATTTTTATAGTTAAATGTCAATTTAGAGATGGGAGAGGGATTA			1910
DB						
QY	1910	TTTTGG	GAAGTTGAATTTTTATAGTTAAATGTCAATTTAGAGATGGGAGAGGGATTA			1969
DB						
QY	1911	TACTGC	AGGAGCTTCAGCCATGTTGTGAAACTGATAAAGCAACTTAGCAAGGCTTCTT			1970
DB						
QY	1970	TACTGC	AGGAGCTTCAGCCATGTTGTGAAACTGATAAAGCAACTTAGCAAGGCTTCTT			2029
DB						
QY	1971	TTCAAT	ATTTTTTTATGTTTTCATTAATAAGTCTTAGGTAACTAGTAGGATAGAAACACTG			2030
DB						
QY	2030	TTCAAT	ATTTTTTTATGTTTTCATTAATAAGTCTTAGGTAACTAGTAGGATAGAAACACTG			2089
DB						
QY	2031	TGTCGG	AGAGTAAAGGAGAGCTACTATTGATTAGACCTTAA	CCAGGTTAACTGCAA		2090
DB						
QY	2090	TGTCGG	AGAGTAAAGGAGAGCTACTATTGATTAGACCTTAA	CCAGGTTAACTGCAA		2149
DB						
QY	2091	GAAGAG	GGGGATACTTTTCAGCTTTCCATGTAACTGTATGCAATAAGCAACTAGTGCCA			2150
DB						
QY	2150	GAAGAG	GGGGATACTTTTCAGCTTTCCATGTAACTGTATGCAATAAGCA			

D	b		2270	TGATGGACAAATAACAGGCCCAAGCCTGTGGTATGATGTGCACACTTGCTAGACTCAGAA	2329
			2271	AAAATACTACTCTCATAAATGGGTGGGAGTATTTTGGTCACAACCTACTTGGCTGGGCTG	2330
			2280	AAAATACTACTCTCATAAATGGGTGGGAGTATTTTGGTCACAACCTACTTGGCTGGGCTG	2389
			2331	AGTGAAGGAATCATATTCCATATATTCATTTATTCATGGACATTTTAGTTAGTGCCTTTTA	2390
			2390	AGTGAAGGAATGATATTCATATATTCATTTATTCATGGACATTTTAGTTAGTGCCTTTTA	2449
			2391	TATACCAGGCATGATCTCGAGTGACACTCTTGGTGTATATTTCCATGGACATTTTGTATAGTCG	2450
			2450	TATACCAGGCATGATCTCGAGTGACACTCTTGGTGTATATTTCCAAATTTTGTACAGTCG	2509
			2451	CTGCACATATTTCAAATCAAATATTAAGAAGTCTCCAAAAATTTGGTCCCTGGCTTTTCA	2510
			2510	CTGCACATATTTCAAATCAAATATTAAGAAGTCTCCAAAAATTTGGTCCCTGGCTTTTCA	2568
			2511	TGGCAACTTGATCAGTAAGGATTTCCCCTCTGTTGGAACTAAAACCATTTACTATATGT	2570
			2569	TGGCAACTTGATCAGTAAGGATTTCCCCTCTGTTGGAACTAAAACCATTTACTATATGT	2628
			2571	TAGACAGACATTTTTTTTTTTTCCCCCTCCCTGAAAAA- AAATGAGGGAAGAGACA	2624
			2629	TAGACATGACATTTCTTTCTCTCTCCCTGAAAAATAAAAGTGTGGGAAGAGACA	2683
RESULT					
LOCUS					
DEFINITION					
ACCESSION					
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS					
TITLE					
JOURNAL					
FEATURES					
source					

## ORIGIN

Query Match	93.4%	Score 2485;	DB 6;	Length 2693;
Best Local Similarity	97.8%;	Pred. No. 0;		
Matches 2576;	Conservative	0;	Mismatches 20;	Indels 39; Gaps 4;
QY	28	AACCTTGGTGCCTGCGTCCGTGAGAAATTCAGCATGGAATGCTCTACTATTTCTGGAT	87	
Db	50	ACCTTGAGTGCCTGCTCGGTGAGAAATTCAGCATGGAATGCTCTACTATTTCTGGAT	109	
QY	88	TTCTGCTCTGCGTGCAGATTGCGACCTTGATGCGGCCAAACGATTTTCATGATGCTCGG	147	
Db	110	TTCTGCTCTGCGTGCAGATTGCGACCTTGATGCGGCCAAACGATTTTCATGATGCTCGG	169	
QY	148	GCAATGAAAGACCTTCTGCTTACATGAGGAGACAATCAATTAATGGCTGGTCTTCTG	207	
Db	170	GCAATGAAAGACCTTCTGCTTACATGAGGAGACAATCAATTAATGGCTGGTCTTCTG	229	
QY	208	ATGAAATGATGGAAATGAAAACTTACCAGTGTGGAAGCGGGAGACATGAGGTGGA	267	
Db	230	ATGAAATGATGGAAATGAAAACTTACCAGTGTGGAAGCGGGAGACATGAGGTGGA	289	

QY 268 AAAAATCCTCGAAGGAGGCGGTGTGAGGCGGTCTTGACAGTGAATCACCAGCCCTCG 327  
Db |||||  
QY 290 AAAAATCCTCGAAGGAGGCGGTGTGAGGCGGTCTTGACAGTGAATCACCAGCCCTCG 349  
Db |||||  
QY 328 TGGGCTCAAAATATAAATTTGCGGTGAACCTGATATTCCTAGATGCAAAAGGAAGATG 387  
Db |||||  
QY 350 TGGGCTCAAAATAAATTTGCGGTGAACCTGATATTCCTAGATGCAAAAGGAAGATG 409  
Db |||||  
QY 388 CCAATGGCAATAGTCTATGAGAAGAACTGCAAGAAATGAGGTGTTTATCTGCTGATC 447  
Db |||||  
QY 410 CCAATGGCAATAGTCTATGAGAAGAACTGCAAGAAATGAGGTGTTTATCTGCTGATC 469  
Db |||||  
QY 448 CATATGTTTAACTGGACAGCATGTTGAGAGACAGTGCAGGCGGAAATGCGACCGGC 507  
Db |||||  
QY 470 CGTATGTTTAACTGGACAGCATGTTGAGAGACAGTGCAGGCGGAAATGCGACCGGC 529  
Db |||||  
QY 508 AAAGCCATCATACGTCCTCCCTGATGGGAAACCTTTCTCACACCCCGGATGGAGAA 567  
Db |||||  
QY 530 AAAGCCATCATACGTCCTCCCTGATGGGAAACCTTTCTCACACCCCGGATGGAGAA 589  
Db |||||  
QY 568 GATGAAATTCATACGTCCTCCACACACTTGGTCAGTATTTCCAGAAATTTGGACGAT 627  
Db |||||  
QY 590 GATGAAATTCATACGTCCTCCACACACTTGGTCAGTATTTCCAGAAATTTGGACGAT 649  
Db |||||  
QY 628 GTTCAGTGAGAGTTTCTGTGAACACAGCCAAATGTGACACTTGGGCTCAACTCATGGAG 687  
Db |||||  
QY 650 GTTCAGTGAGAGTTTCTGTGAACACAGCCAAATGTGACACTTGGGCTCAACTCATGGAG 709  
Db |||||  
QY 688 TGACTGTCTACAGAGACATGACCGGCGATATGTTCCCATCGCAAGTGAAGATGTGT 747  
Db |||||  
QY 710 TGACTGTCTACAGAGACATGACCGGCGATATGTTCCCATCGCAAGTGAAGATGTGT 769  
Db |||||  
QY 748 ACGTGGTAAACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAAACAGATCGAAAT 807  
Db |||||  
QY 770 ACGTGGTAAACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAAACAGATCGAAAT 829  
Db |||||  
QY 808 CATCCGAGCAACCTTCC - CAAAGATCTCCCATATGTTGATGTCCTGATTCATGATC 866  
Db |||||  
QY 830 CATCCGAGCAACCTTCCCAAAGATCTCCCATATGTTGATGTCCTGATTCATGATC 889  
Db |||||  
QY 867 CTAGCCACTTCTCAATATTTCTACCAATTAATACTCAAGTGAGCTTCGGGATTAATCTG 926  
Db |||||  
QY 890 CTAGCCACTTCTCAATATTTCTACCAATTAATACTCAAGTGAGCTTCGGGATTAATCTG 949  
Db |||||  
QY 927 GCCTGTTTGTTCACCAATCATACTGTGAATCACAGTATGCTCAATGGAACTTCA 986  
Db |||||  
QY 950 GCCTGTTTGTTCACCAATCATACTGTGAATCACAGTATGCTCAATGGAACTTCA 1009  
Db |||||  
QY 987 GCCTTAACCTCATGTGAAAGCTGCAGCACAGGACCTTGTCCGCCACCGCCACCCAC 1046  
Db |||||  
QY 1010 GCCTTAACCTCATGTGAAAGCTGCAGCACAGGACCTTGTCCGCCACCGCCACCCAC 1069  
Db |||||  
QY 1047 CCAGACCTTCAAAACCCACCCCTCTTT----- 1074  
Db |||||  
QY 1070 CCAGACCTTCAAAACCCACCCCTCTTTTAGCAACTCTTAAATCTTATGATTCAAACA 1129  
Db |||||  
QY 1075 ---AGGACCTGTGTTGACAAACCCCTGAGCTGAGTATTCCTGATGAACCTGCC 1130  
Db |||||  
QY 1130 CCCAGGACCTCATGTGTGACAAACCCCTGAGCTGAGTAGGATTCCTGATGAACCTGCC 1189  
Db |||||  
QY 1131 AGATTAACAGATATGGCCACTTTCAAGCCACCATCAATTTGAGAGGAAATCTTAGAGG 1190  
Db |||||  
QY 1190 AGATTAACAGATATGGCCACTTTCAAGCCACCATCAATTTGAGAGGAAATCTTAGAGG 1249  
Db |||||  
QY 1191 TTAACATCATCCAGATGACAGAGCTCTGATGCCGTGCCATGGCCTGAAGCTCCCTAA 1250  
Db |||||  
QY 1250 TTAACATCATCCAGATGACAGAGCTCTGATGCCGTGCCATGGCCTGAAGCTCCCTAA 1309  
Db |||||  
QY 1251 TAGACTTTGTGCTGACCTGCCAGGAGGATTCACCGAGGTCTGTACCATCATTTGTG 1310  
Db |||||  
QY 1310 TAGACTTTGTGCTGACCTGCCAGGAGGATTCACCGAGGTCTGTACCATCATTTCTG 1369  
Db |||||

QY 1311 ACCCACTCTCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGT 1370  
Db |||||  
QY 1370 ACCCACTCTCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGT 1429  
Db |||||  
QY 1371 GTCTGCTGACTGTGAGACGAACCTTCAATGGGCTCTGGGAGCTACTGTGTGAACCTCACCC 1430  
Db |||||  
QY 1430 GTCTGCTGACTGTGAGACGAACCTTCAATGGGCTCTGGGAGCTACTGTGTGAACCTCACCC 1489  
Db |||||  
QY 1431 TGGGGATGACACAAAGCTCTGCTCACGAGACCTGATTTCTGTTCTCTGACAGAGACC 1490  
Db |||||  
QY 1490 TGGGGATGACACAAAGCTCTGCTCACGAGACCTGATTTCTGTTCTCTGACAGAGACC 1549  
Db |||||  
QY 1491 CAGCTCTGCTTTAAAGATGGCAAAACAGTGCCTGATCTCCGTTGCTGCTGGCCATAT 1550  
Db |||||  
QY 1550 CAGCTCTGCTTTAAAGATGGCAAAACAGTGCCTGATCTCCGTTGCTGCTGGCCATAT 1609  
Db |||||  
QY 1551 TTGCTACTGTGATCTCCCTCTTGGTGTACAAAACACAAAGGAATACAACCCCAATAGAAA 1610  
Db |||||  
QY 1610 TTGCTACTGTGATCTCCCTCTTGGTGTACAAAACACAAAGGAATACAACCCCAATAGAAA 1669  
Db |||||  
QY 1611 ATAGTCTCTGGGAATGTGTGAGAAAGGCTGAGTGTCTTTCTCAACCTGCAAAAAG 1670  
Db |||||  
QY 1670 ATAGTCTCTGGGAATGTGTGAGAAAGGCTGAGTGTCTTTCTCAACCTGCAAAAAG 1729  
Db |||||  
QY 1671 CCGTGTCTTCCCGGAAACAGGAAAGGATCCGCTACTCAAAAACCAAGAAATTTAAAG 1730  
Db |||||  
QY 1730 CCGTGTCTTCCCGGAAACAGGAAAGGATCCGCTACTCAAAAACCAAGAAATTTAAAG 1789  
Db |||||  
QY 1731 GAGTTTCTTAAATTTGCACTGTGTTCTGAAAGCTCACCTTTTCAGTGCCTATGATGTGAGA 1790  
Db |||||  
QY 1790 GAGTTTCTTAAATTTGCACTGTGTTCTGAAAGCTCACCTTTTCAGTGCCTATGATGTGAGA 1849  
Db |||||  
QY 1791 TGTGCTGAGTGGCTATTTAACTTTTCTTAAAGATTTTGTGTTAAATAGATATTTGFG 1850  
Db |||||  
QY 1850 TGTGCTGAGTGGCTATTTAACTTTTCTTAAAGATTTTGTGTTAAATAGATATTTGFG 1909  
Db |||||  
QY 1851 TTTCCGGAAGTTGAATTTTATAGTTTAAATGCTATTTAGAGATGGGAGAGGATTA 1910  
Db |||||  
QY 1910 TTTCCGGAAGTTGAATTTTATAGTTTAAATGCTATTTAGAGATGGGAGAGGATTA 1969  
Db |||||  
QY 1911 TACTGCAAGGAGCTTCAGCCATGTTGTAACCTGATAAAAGCACTTAGCAAGGCTCTTT 1970  
Db |||||  
QY 1970 TACTGCAAGGAGCTTCAGCCATGTTGTAACCTGATAAAAGCACTTAGCAAGGCTCTTT 2029  
Db |||||  
QY 1971 TTTCAATTTTATGTTTTCATCTATAGTCTTAGTAACTAGTAGGATAGAAACTG 2030  
Db |||||  
QY 2030 TTTCAATTTTATGTTTTCATCTATAGTCTTAGTAACTAGTAGGATAGAAACTG 2089  
Db |||||  
QY 2031 TGTCCCGAGAGTAAAGGAGAGAGCTACTATTTAGTTAGAGCTAACCCAGGTTAACTGCAA 2090  
Db |||||  
QY 2090 TGTCCCGAGAGTAAAGGAGAGAGCTACTATTTAGTTAGAGCTAACCCAGGTTAACTGCAA 2149  
Db |||||  
QY 2091 GAAGAGGCGGATACCTTCAGCTTTCCATGTAACTGTATGCAATAAGCCAACTAGTCCA 2150  
Db |||||  
QY 2150 GAAGAGGCGGATACCTTCAGCTTTCCATGTAACTGTATGCAATAAGCCAACTAGTCCA 2209  
Db |||||  
QY 2151 GTTTCTTAAGATCATGTTTCCAAAGCTAACTAAATCCCACTTCAATACACTCATGAACTCC 2210  
Db |||||  
QY 2210 GTTTCTTAAGATCATGTTTCCAAAGCTAACTAAATCCCACTTCAATACACTCATGAACTCC 2269  
Db |||||  
QY 2211 TGATGAAACAATAACAGGCGCCAGGCTGTGTTGATGTGACACCTGCTAGACTCAGAA 2270  
Db |||||  
QY 2270 TGATGAAACAATAACAGGCGCCAGGCTGTGTTGATGTGACACCTGCTAGACTCAGAA 2329  
Db |||||  
QY 2271 AAAATACTACTCTCAATAATGGGTGGAGTATTTTGGTGACAACTACTTTGCTGGCTG 2330  
Db |||||  
QY 2330 AAAATACTACTCTCAATAATGGGTGGAGTATTTTGGTGACAACTACTTTGCTGGCTG 2389  
Db |||||  
QY 2331 AGTGAAGGAATGATATTCATATTTTCCATGGAATTTAGTTAGTGTGCTTTT 2390  
Db |||||  
QY 2390 AGTGAAGGAATGATATTCATATTTTCCATGGAATTTAGTTAGTGTGCTTTT 2449  
Db |||||  
QY 2391 TATACAGGAGATGCTGTGAGTGACACTCTTGTGTATATTTCCAAATTTTGTATAGTGC 2450  
Db |||||



```
QY 1047 CCAGACCTTCAAAACCCACCCCTTCTTT----- 1074
Db 1070 CCAGACCTTCAAAACCCACCCCTTCTTTAGCAACTACTCTAAATCTTATGATTCAAACA 1129
QY 1075 ----AGGACCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCC 1130
Db 1130 CCCAGGACCTTACTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCC 1189
QY 1131 AGATTAAAGATAGGCGACACTTTCAAGCCACCATCAAAATTTGTAGAGGAATCTTAGAGG 1190
Db 1190 AGATTAAAGATAGGCGACACTTTCAAGCCACCATCAAAATTTGTAGAGGAATCTTAGAGG 1249
QY 1191 TTAACATCATCCAGATGACAGACGTCCTGATGCGGTGCCATGCCCTGAAAGCTCCCTAA 1250
Db 1250 TTAACATCATCCAGATGACAGACGTCCTGATGCGGTGCCATGCCCTGAAAGCTCCCTAA 1309
QY 1251 TAGACTTTGTGCTGACCTGCCAAGGAGCAATCCACGAGGTCTGTACCATCAATTTCTG 1310
Db 1310 TAGACTTTGTGCTGACCTGCCAAGGAGCAATCCACGAGGTCTGTACCATCAATTTCTG 1369
QY 1311 ACCCCACCTGGAGATCACCCAGAACACAGTCTGCGAGCTACTGTGTGAACCTCACCC 1370
Db 1370 ACCCCACCTGGAGATCACCCAGAACACAGTCTGCGAGCTACTGTGTGAACCTCACCC 1429
QY 1371 GTCTGCTGACTGTGAGACGAACTTCAATGGGTCTGGAGCTACTGTGTGAACCTCACCC 1430
Db 1430 GTCTGCTGACTGTGAGACGAACTTCAATGGGTCTGGAGCTACTGTGTGAACCTCACCC 1489
QY 1431 TGGGGATGACACAAGCCTGCTCTACGAGCACCTGATTTCTGTTCTCTGACAGACCC 1490
Db 1490 TGGGGATGACACAAGCCTGCTCTACGAGCACCTGATTTCTGTTCTCTGACAGACCC 1549
QY 1491 CAGCTCGCTTTAAGATGGCAACAGTCCCTGATCTCGCTGGCTGCTGCTGCCATAT 1550
Db 1550 CAGCTCGCTTTAAGATGGCAACAGTCCCTGATCTCGCTGGCTGCTGCTGCCATAT 1609
QY 1551 TTGTCACGTGATCTCCCTCTTGGTGACAAAAACACAAAGGAATACAAACCCATGAAA 1610
Db 1610 TTGTCACGTGATCTCCCTCTTGGTGACAAAAACACAAAGGAATACAAACCCATGAAA 1669
QY 1611 ATAGTCTGGGAATGGTGAGAACGAAAGCCTGAGTGCTTTCTCAACCGTGCAAAAG 1670
Db 1670 ATAGTCTGGGAATGGTGAGAACGAAAGCCTGAGTGCTTTCTCAACCGTGCAAAAG 1729
QY 1671 CCGTGCTCTCCGCGAACCAGAAAGGATCGCTACTCAAAACCAAGAAATTTAAAG 1730
Db 1730 CCGTGCTCTCCGCGAACCAGAAAGGATCGCTACTCAAAACCAAGAAATTTAAAG 1789
QY 1731 GAGTTTCTTAAATTTGACCTTCTTGAAGCTCACTTTTCAGTGCCATGATGTGAGA 1790
Db 1790 GAGTTTCTTAAATTTGACCTTCTTGAAGCTCACTTTTCAGTGCCATGATGTGAGA 1849
QY 1791 TGTGCTGGAGTGCTAATAACCTTTTCTTAAAGATTTATGTTAAATAGATATTGTGG 1850
Db 1850 TGTGCTGGAGTGCTAATAACCTTTTCTTAAAGATTTATGTTAAATAGATATTGTGG 1909
QY 1851 TTTGGGGAAGTTGAATTTTATAGTTTAAATGCTCAATTTAGAGATGGGAGGGATTA 1910
Db 1910 TTTGGGGAAGTTGAATTTTATAGTTTAAATGCTCAATTTAGAGATGGGAGGGATTA 1969
QY 1911 TACTGAGGACGCTTCAGGCATGTTGTGAAACTGATAAAGCAACTTACAGGCTTCCT 1970
Db 1970 TACTGAGGACGCTTCAGGCATGTTGTGAAACTGATAAAGCAACTTACAGGCTTCCT 2029
QY 1971 TTCATTTATTTTATGTTTCACTTATAAAGCTTTAGTAACTAGTAGGATGAAACACTG 2030
Db 2030 TTCATTTATTTTATGTTTCACTTATAAAGCTTTAGTAACTAGTAGGATGAAACACTG 2089
QY 2031 TGTCCCGAGATAGGAGAGACTACTTATGATTAGAGCCTAACCCAGGTTAACTGCAA 2090
Db 2090 TGTCCCGAGATAGGAGAGACTACTTATGATTAGAGCCTAACCCAGGTTAACTGCAA 2149
```

```
QY 2091 GAAGAGCGGATACTTTTCAGCTTTCATGTAACTGTATGATAAAGCCAATGTAGTCCA 2150
Db 2150 GAAGAGCGGATACTTTTCAGCTTTCATGTAACTGTATGATAAAGCCAATGTAGTCCA 2209
QY 2151 GTTTCTAAGATCANGTTCCTCAAGCTAACTGAATCCCACTTCAATACACACTCATGAACCTC 2210
Db 2210 GTTTCTAAGATCANGTTCCTCAAGCTAACTGAATCCCACTTCAATACACACTCATGAACCTC 2269
QY 2211 TGATGGAACAATAACAGGCCCAAGCCTGTGTGTATGATGTGCACACTTCTCTAGACTCAGAA 2270
Db 2270 TGATGGAACAATAACAGGCCCAAGCCTGTGTGTATGATGTGCACACTTCTCTAGACTCAGAA 2329
QY 2271 AAAATACTACTCTCATAAATGGGTGGAGTATTTTGGTGACCAACTACTTTCCTTGGCTG 2330
Db 2330 AAAATACTACTCTCATAAATGGGTGGAGTATTTTGGTGACCAACTACTTTCCTTGGCTG 2389
QY 2331 AGTGAAGGAATGATTAATTCATATTAATTTATTCATGATGATGATGATGATGATGATGAT 2390
Db 2390 AGTGAAGGAATGATTAATTCATATTAATTTATTCATGATGATGATGATGATGATGATGAT 2449
QY 2391 TATACCAAGGATGATGCTGAGTGACACTCTTGTGTATATTTCCAAATTTTGTATAGTCG 2450
Db 2450 TATACCAAGGATGATGCTGAGTGACACTCTTGTGTATATTTCCAAATTTTGTATAGTCG 2509
QY 2451 CTGCACATTTTGAATCAAAAATTAATTAAGACTTTTCCAAATTTTGGTCCCTGTTTTTCA 2510
Db 2510 CTGCACATTTTGAATCAAAAATTAATTAAGACTTTTCCAAATTTTGGTCCCTGTTTTTCA 2568
QY 2511 TGGCAACTTTCATCAGTAAGGATTTCCCTCTGTTTGGAACTTAAACCACTTATATATGT 2570
Db 2569 TGGCAACTTTCATCAGTAAGGATTTCCCTCTGTTTGGAACTTAAACCACTTATATATGT 2628
QY 2571 TAGACAAGACATTTTTTTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2624
Db 2629 TAGACATGACATTTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2683
```

## RESULT 10

```
AX677738
LOCUS AX677738 1683 bp DNA linear PAT 27-MAR-2003
DEFINITION Sequence 516 from Patent WO02086122.
ACCESSION AX677738
VERSION AX677738.1 GI:29335143
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Legrain,P. and Daviet,L.
TITLE Protein-protein interactions in adipocytes
JOURNAL Patent: WO 02086122-A 516 31-OCT-2002;
Hybridomics (FR)
FEATURES
source
1. .1683
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
```

## ORIGIN

```
Query Match 62.8%; Score 1671; DB 6; Length 1683;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1682; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 60 ATGGAATGCTCTACTATTTCTCTGGGATTTCTGCTCTGGCTGCAAGATTGCCACTTGAT 119
Db 1 ATGGAATGCTCTACTATTTCTCTGGGATTTCTGCTCTGGCTGCAAGATTGCCACTTGAT 60
QY 120 GCCGCAAAACGATTTTCATGATGTCTGGCAATGAAGACCTTCTGCTTACATGAGGAG 179
Db 61 GCCGCAAAACGATTTTCATGATGTCTGGCAATGAAGACCTTCTGCTTACATGAGGAG 120
QY 180 CACAATCAATTAATGGCTGTGCTTCTGATGAAATGATGGAATGAAAAACTCTACCCA 239
```



Db	121	CACAAATCAATTAATGGCTGGTCTTCTGATGAAATGACTGGAAATGAAAAAATCTTACCCCA	180
Qy	240	GTGTGGAAGCGGGAGACATGAGGTGGAAAAAACTCCTGGAAGGGAGGGCGTGTGCAGGCG	299
Db	181	GTGTGGAAGCGGGAGACATGAGGTGGAAAAAATCCTCTGGAAAGGGAGGGCGTGTGCAGGCG	240
Qy	300	GTCTCAGCAGTGCACATCACAGCCCTCGTGGGTCAAATATATAACATTTGCGGTGAACCTG	359
Db	241	GTCTCGACCAAGTGCATCACAGCCCTCTGTGGGTCAAATATATAACATTTGCGGTGAACCTG	300
Qy	360	ATATTCCCTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAAGTGC	419
Db	301	ATATTCCCTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAAGTGC	360
Qy	420	AGAAATGAGCTGGTTTATCTGCTGATCCATATGTTTTTCAAACTGGACAGCATGGTTCAGAG	479
Db	361	AGAAATGAGCTGGTTTATCTGCTGATCCATATGTTTTTCAAACTGGACAGCATGGTTCAGAG	420
Qy	480	GACAGTCACGGGGAATAATGCCCGCCAAAGCCATCAATACGTCTTCCCTGATGGGAAA	539
Db	421	GACAGTCACGGGGAATAATGCCCGCCAAAGCCATCAATACGTCTTCCCTGATGGGAAA	480
Qy	540	CCTTTTCCTCACCAACCCCGGATGGAGAAGATGGAATTTTCACTPACGTCTTCCACACATTT	599
Db	481	CCTTTTCCTCACCAACCCCGGATGGAGAAGATGGAATTTTCACTPACGTCTTCCACACATTT	540
Qy	600	GGTCAGTATTTTCAGAAATTTGGGACGATGTTCTGATGAGTGTTCCTGTGAACACAGCCAAT	659
Db	541	GGTCAGTATTTTCAGAAATTTGGGACGATGTTCTGATGAGTGTTCCTGTGAACACAGCCAAT	600
Qy	660	GTGACACTTGGGCTCAACTCATGGAAGTGAATTTCTACAGAAAGATGGAAGTGGACGGGCATAT	719
Db	601	GTGACACTTGGGCTCAACTCATGGAAGTGAATTTCTACAGAAAGATGGAAGTGGACGGGCATAT	660
Qy	720	GTTCCCATTCGCACAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCCTGTGTTTGTG	779
Db	661	GTTCCCATTCGCACAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCCTGTGTTTGTG	720
Qy	780	ACTATGTCGAGAAGACGATCGAAATTCATCCGACGAACCTTCC-CANAGATCTCCCC	838
Db	721	ACTATGTCGAGAAGACGATCGAAATTCATCCGACGAACCTTCC-CANAGATCTCCCC	780
Qy	839	ATTATGTTTGATGTCCTGATTCATGATCCCTAGCCACTTCTCAATTAATCTACCATTAAC	898
Db	781	ATTATGTTTGATGTCCTGATTCATGATCCCTAGCCACTTCTCAATTAATCTACCATTAAC	840
Qy	899	TACAAGTGGAGCTTCGGGGATAATACTGGCTGTGTTTTCACCAATCATACTGTGAAT	958
Db	841	TACAAGTGGAGCTTCGGGGATAATACTGGCTGTGTTTTCACCAATCATACTGTGAAT	900
Qy	959	CACAGTATGTGCTCAATGGAACCTTCAGCCCTTAACTCACTGTGAAGCTGCAGCACCA	1018
Db	901	CACAGTATGTGCTCAATGGAACCTTCAGCCCTTAACTCACTGTGAAGCTGCAGCACCA	960
Qy	1019	GGACCTTGTCCGCACAGCCACACCCAGACCTTCAAAACCCACCCCTCTTTAGGA	1078
Db	961	GGACCTTGTCCGCACAGCCACACCCAGACCTTCAAAACCCACCCCTCTTTAGGA	1020
Qy	1079	CCTGCTGTGTGAACCCCTCTGGAGCTGATGAGTATTCCTGATGAAAACTGCAGATTAAC	1138
Db	1021	CCTGCTGTGTGAACCCCTCTGGAGCTGATGAGTATTCCTGATGAAAACTGCAGATTAAC	1080
Qy	1139	AGATATGGCCATTTTCAGCCACCATCATCAATTTGTAGAGGGAATCTTAGAGGTTAACATC	1198
Db	1081	AGATATGGCCATTTTCAGCCACCATCATCAATTTGTAGAGGGAATCTTAGAGGTTAACATC	1140
Qy	1199	ATCCAGATCAGACAGCTCTCTGATGCGGTGCGCATGGCTGAAAGCTCCCTAATAGACTTT	1258
Db	1141	ATCCAGATCAGACAGCTCTCTGATGCGGTGCGCATGGCTGAAAGCTCCCTAATAGACTTT	1200
Qy	1259	GTCTGACCTGCCAAGGGAGCATTCCTCAAGAGGTCTGTACATCATTTCTGACCCACCC	1318

Db	1201	GTCTGACCTGCACAAAGGAGCATTC	CCACGAGGCTCTGTACCATCATTTCTGACCCACC	1266
Qy	1319	TGGGAGATCACCCAGAACACAGTCT	CAGCCCTGTGGATGTGATGAGATGTGTCTGCTG	1378
Db	1261	TGGGAGATCACCCAGAACACAGTCT	CAGCCCTGTGGATGTGATGAGATGTGTCTGCTG	1320
Qy	1379	ACTGTGAGAGCAACCTTCAATGGT	CTGGAGCTGAGAGTCTGAGACCTCAACCTGGGGGAT	1438
Db	1321	ACTGTGAGAGCAACCTTCAATGGT	CTGGAGCTGAGAGTCTGAGACCTCAACCTGGGGGAT	1380
Qy	1439	GACACAGCCTGGCTCTCACGAGC	CACCTCATTTCTGCTCCTCACGAGACCCAGCCCTCG	1498
Db	1381	GACACAGCCTGGCTCTCACGAGC	CACCTCATTTCTGCTCCTCACGAGACCCAGCCCTCG	1440
Qy	1499	CCTTTAAGGATGGCAACACGATG	CCCTGGCTGTGGTCTGTGGCCATATTGTGCACT	1558
Db	1441	CCTTTAAGGATGGCAACACGATG	CCCTGGCTGTGGTCTGTGGCCATATTGTGCACT	1500
Qy	1559	GTGATCTCCCTCTGTGTACAA	AAACACAAAGGAATACACCAATAGAAAAATAGTCT	1618
Db	1501	GTGATCTCCCTCTGTGTGTAC	AAAAACACAAAGGAATACACCAATAGAAAAATAGTCT	1560
Qy	1619	GGGAATCTGGTTCAGAAAG	CGCTCAGTGTCTTTCTCAACCGTCAAAAGCCGTGTC	1678
Db	1561	GGGAATCTGGTTCAGAAAG	CGCTCAGTGTCTTTCTCAACCGTCAAAAGCCGTGTC	1620
Qy	1679	TTCCCGGGAAACCGAGAA	AGGATCCGCTACTCAAAAACCCAGAAATTTAAGAGATTTCT	1738
Db	1621	TTCCCGGGAAACCGAGAA	AGGATCCGCTACTCAAAAACCCAGAAATTTAAGAGATTTCT	1680
Qy	1739	TAA	1741	
Db	1681	TAA	1683	
RESULT 11				
BC061725				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				



AUTHORS	TITLE	REMARK	FEATURES
Strausberg, R. Direct Submission Submitted (05-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA NH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a> Contact: MGC help desk Email: <a href="mailto:cgabbs-r@mail.nih.gov">cgabbs-r@mail.nih.gov</a> Tissue Procurement: Jeff Green/Paturu Kondaiah, NCI. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: <a href="http://www.nisc.nih.gov/">http://www.nisc.nih.gov/</a> Contact: <a href="mailto:nisc_mgc@nhgri.nih.gov">nisc_mgc@nhgri.nih.gov</a> Akhter, N., Ayèle, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S., Dietsch, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Series: IRAK Plate: 137 Row: c Column: 23 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 18959233. Location/Qualifiers		

## source

source	1. .2282 /organism="Rattus norvegicus" /mol_type="mRNA" /db_xref="taxon:10116" /clone="MGC:72320 IMAGE:5598392" /tissue_type="Prostate, pool of NCI_CGAP_Pr30, 40, 41 and NCI_CGAP_Pr29, 39, 42" /clone_lib="NCI_CGAP_Pr49" /lab_host="DH10B" /notes="Vector: pCMV-SPORT6.1" 1. .2282 /genes="Gpnmb" /db_xref="LocusID:113955" /db_xref="RGD:71008" 58. .1776 /codon_start=1 /product="glycoprotein (transmembrane) nmb" /protein_id="AAH61725.1" /db_xref="GI:38197644" /db_xref="LocusID:113955" /translation="MESLQGLVFLALLAAGLPLQAAKFRDVLGHQYQDDHMRNNQL RGMSDENWDEQLYVWRGEGRWKSGEGVQAALTSDSPALVGSNTTFVNILVF PQCREDANGNIYERNCSDELAESPYYNMTGADDEWEDSTSQGHLPFPDGK PPFPRGKRGMNFVYHTLQGYQKLGRCARSNTVNLTVGPOMEVIVFRHGR AIPKISKVDVYITQIPIFVTMYOKNDRNSDETEFLRDLPIFFDLVLDPSHFLNY SALSYPWFGDNTGLFVSNHMLNHTYLVNGTFNFLTQVTAVPGPCPSTTPSPSSIT SPSPSPSPSTLSPSPSLMPTGKMSWELSDISNENCRINRYGFATITIVDGLLEV NIIQAVDPIPTQPDSLMDFITCKGATPEACTIISDPTCQIAQNRVCSPPVADE LCLLSYRRAFNSGTYCVNFTLGDGDSALTALISPGKDLGSPURTNGVLISIG LAMFVFMVTWITLLYKKHKTYKPIGNCITRNVVKGSLSVFLSHAKAPSRGDRKDPILQ DKFPMWL" 844. .1014 /notes="PKD; Region: Repeats in polycystic kidney disease 1 (PKD1) and other proteins" /db_xref="CDD:smart00089"
gene	
CDS	
misc_feature	

## ORIGIN

Query Match 39.5%; Score 1051.2; DB 10; Length 2382;  
Best Local Similarity 74.1%; Pred. No. 3.8e-248;  
Matches 1405; Conservative 0; Mismatches 443; Indels 48; Gaps 4;

```
Db 1109 CCTCTTTAATGCTACTGGCCCAAAATCAATGAGCTGAGTGCATTTCCAAATGAACACT 1168
Qy 1128 GCCAGATTAAACAGATATGGCCACTTTCAAGCCACCAATCAAAATGTAGAGGGAATCTTAG 1187
Db 1169 GCCGAATAAACAGATATGGTTACTTCAGAGCCACCAATCAAAATGTAGATGAATCTTAG 1228
Qy 1188 AGTTTAACATATCCAGATACAGACGCTCTGATCGCGTGGCCATGGCCCTGAAAGCTCCC 1247
Db 1229 AAGTCAACATATCCAGTACAGATGTCCCAATGCCCAACACCGCAGCCCTGCAACATCAC 1288
Qy 1248 TAATAGACTTTGTGCTGACCTGCCAAGGAGACATTCACCGAGAGCTGTGTACCATCATTT 1307
Db 1289 TGAATGACTTATGTGACCTGTGCAAGGGGCCACTCCACGGAAGCTGTACGATCATCT 1348
Qy 1308 CTGACCCCACTGGAGATACCCAGAACACAGTGTGACGCCCTGTGGATGATGAGA 1367
Db 1349 CTGACCCCACTGCCAGATGCCCAAGAAAGGGTGTGACGCCCGGTGGCTGTGGATGAGC 1408
Qy 1368 TGTGTCTGCTGACTGTGAGAGAACCTTCAATGGGTCTGGAGCTACTGTGTGAACCTCA 1427
Db 1409 TGTGCTCTCTGTCGAGGAGACCTTCAATGGGTCCGGCAGCTACTGTGTGAATTTCA 1468
Qy 1428 CCTGGGGGATGACACAAAGCTGCTCTCAGAGACCCCTGATTTCTGTCTCTGACAGAG 1487
Db 1469 CTCTGGAGAGAGTGAAGCTGCTCCTCACCAGGCCCTGATCTCTATCCCTGGCAAG 1528
Qy 1488 ACCAGCCTCGCTTTAAGGATGCAACAGTGTGCTGATCTCGGTGGCTGTGGCCCA 1547
Db 1529 ACCTAGGCTCCCTCTGAGAACAGTGAATGGTGTCTGATCTCCATTTGGCTGTGGCCCA 1588
Qy 1548 TATTGTCACTGTGATCTCCCTCTTGTGTACAAAACAAAGAAATACAAACCAATAG 1607
Db 1589 TGTTGTCAACCATGGTTACCATCTTGTGTACAAAACAAAGAAATACAAACCAATAG 1648
Qy 1608 AAAATAGTCTCGGAATGTGTGAGAAAGAGCTGAGTGTCTTCTTCAACCGTGCAA 1667
Db 1649 GAAACTGCACAGGAACTGTGTGAGGCAAGGCTGAGTGTCTTCTCAGCCATGAA 1708
Qy 1668 AAGCCGTGTCTTCCCGGAAACACAGAAAGATCGCTACTCAAAACCAAGAAATTA 1727
Db 1709 AAGCCCGTCTTCCCGGAGAGACCGGAGAGAGATCCACTGCTCCAGGACAAAGCATGGA 1768
Qy 1728 AAGAGTGTCTTAAATTCGACCTGTCTTGTGAGCTCACTTTTCAAGTGCATTTGATGTG 1787
Db 1769 TGCTCTAAGTCTTCACTCTCTGACTGGAAACCCACTCTTCTGTG-CATGTAATGTG 1827
Qy 1788 AGATGTCTGGAG---TGGCTATTAACTTTTCTTCTTAAAGATATTGTTAAATAGAT 1843
Db 1828 AGCTGTGAGAAATACATGACTGGTAGCTGTGTTTCTACGGATATTGTTAAATGTAT 1887
Qy 1844 ATTGTGTTTGGGGAAGTTGAATTTTATAGTTA 1879
Db 1888 ATCATGGTTAGGAGCGTAGTTAATGGCAATTTA 1923

RESULT 12
BD062749 2303 bp DNA linear PAT 27-AUG-2002
LOCUS Modulators of tissue regeneration.
DEFINITION
ACCESSION BD062749
VERSION
KEYWORDS JP 2001505761-A/3.
SOURCE Saccharomyces cerevisiae (baker's yeast)
ORGANISM Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
REFERENCE
1 (bases 1 to 2303)
AUTHORS Nadel,M.S., Bonventre,J.V., Hession,C.A., Ichimura,T., Wei,H. and
Cate,R.L.
TITLE Modulators of tissue regeneration
JOURNAL Patent: JP 2001505761-A 3 08-MAY-2001;
BIOGEN INC
```

```
COMMENT PN JP 2001505761-A/3
PD 08-MAY-2001
PF 23-MAY-1997 JP 1997542986
PR 24-MAY-1996 US 60/018228,23-AUG-1996 US 60/023442 PI
MICHELE SANICOLA NADEL, JOSEPH V BONVENTRE, CATHERINE A HESSION, PI
TAKAHARU ICHIMURA, HENRY WEI, RICHARD L CATE
PC C12N15/12,C07K14/47,C12N15/62,C07K16/18,A61K38/16,G01N33/50,
C12Q1/68,
PC C12N1/21,C12N5/10,C12N5/12,A61K48/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
FEATURES
source 1..2303
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"
ORIGIN
Query Match 39.5%; Score 1051.2; DB 6; Length 2303;
Best Local Similarity 74.1%; Pred. No. 3.8e-248;
Matches 1405; Conservative 0; Mismatches 443; Indels 48; Gaps 4;
Qy 31 CTGTGGTCTGCTGCTGCTGAGAAATTCAGCATGGAATCTCTACTATTTCCTGGGATTC 90
Db 78 CTGGCTGAGGGCGGGGCTCCGAGTCAGCATGGAAGTCTCTGGGGTCTCTGGTATTTC 137
Qy 91 TGCTCTGGCTGCAAGATTGCCACTTCATGCCGCCAAACGATTTTCATGATGTCTGGGCA 150
Db 138 TGCTGTGGCTGAGGACTCCGCTCCAGCGGGCAAGCGGTTCCGTGATGTCTGGGCC 197
Qy 151 ATGAAGACCTTCTGCTTACATGAGGAGCACCAATCAATTAATTAATGGCTGCTTCTGATG 210
Db 198 ATGACAGTATCCGATCACAATGAGGAGAACCAATTAACGTGGCTGCTTCTCAGATG 257
Qy 211 AAAATGATGGAATGAAACCTTACCCAGTGTGGAAGCGGGAGACATGAGGTGGAATA 270
Db 258 AAAATGAATGGATGAACAGCTGATCCAGTGTGGAGAGGGAGAGGGAGATGGAAG 317
Qy 271 ACTCTGGAAGGGAGGGCGGTGTCAGGGGTCTGACCGAGTCTCACCAGCTCCTCGTGG 330
Db 318 ACTCTCGGAGAGGGCGGTGTGAGGAGCGCCCTTAACAGTGTATCCGCGCTTGGTGG 377
Qy 331 GCTCAAAATATAAATTTGCGGTGAACCTGATATTCCTTAGATGCCAAAGGAAGATGCCA 390
Db 378 GTTCAATATCACCTTGTAGTGAACCTGGTGTCCCAGATGCCAGAGGAAGATGCCA 437
Qy 391 ATGCCAATAGTCTATGAGAAAGAACTGCAAGAAATGAGGCTGGTTTATCTGCTGATCAT 450
Db 438 ACGCAATATCTCTATGAGAGGAATGCAAGTGTATTTGGAGCTGGCTTCTGACCCGT 497
Qy 451 ATGTTTACAACTGCACAGCATGCTCAGAGGACAGTGCAGGGGAAAATCGCACCGGCAAA 510
Db 498 ATGTCTACAATGGACCAACAGCGGCAGACGATGAGGACTGGGAAGACAGCACCCAG 557
Qy 511 GCCATCATACGTCTTCCCTGATGGGAAACCTTTTCTTCTCAGCACCCCCGGATGGGAAGAT 570
Db 558 GCCAGACCTCAGTTCCTCCGACGGGAAGCCCTTCCCTCGCCCCACGACGGAAGAAAT 617
Qy 571 GGAATTTTCATCTAGCTTTCCACACACTGTGTCAGTATTTCAGAAATTTGGACGATGTT 630
Db 618 GGAACCTTCTAGCTTTCCACACACTGTGTCAGTATTTCAGAAAGCTGGGTGGTGT 677
Qy 631 CAGTGAGATTTCTGTGAACACAGCAATGTGACACTTTGGGCTCAACTCATCTGGAAGTGA 690
Db 678 CAGCACGAGTTTCTATAAACACAGTCAACTTGACAGTTTGGGCTCAGGTCTATGGAATGA 737
Qy 691 CTGTCTACAGAGACATGAGCGGGCATATGTTCCCATCGCAACAGTGAAGATGTGTACG 750
Db 738 TTGCTTTTCGAAGACACAGCGCGGCATACATTTCCCATCTCCAAAGTGAAGACGTGTATG 797
Qy 751 TGGTAACAGATCAGATTCTCTGTGTTGTGACTATGTTCCAGAGAACGATCGAAATTCAT 810
```



14 CCTGACTGGTTGCGAGCGCTCGAGTCAGCATGGAAGATCTCTCGGGGTCCTGGGATT 73  
90 CTGCTCTGCTGCGAGATTGCCACTTGTGATGCGCGCAAAAGATTTCAATGATGCTGGGC 149  
74 CTGCTGCTGCTGCGAGACTGCTCTCCAGGCTGCCAAGCGATTTGCTGATGCTGGGC 133  
150 AATGAAGACCTTCTGCTTACATGAGGGAGCACAAATCAATTAATGCTGCTTCTGAT 209  
134 CATGAACAGTATCCCAATCATATGAGAGAGCAACACCAATTAATGCTGCTTCTGAT 193  
210 GAAATGACTGGAATGAAATCTACCCAGTGTGGAAGCGGGAGAGACATGAGGTGAAA 269  
194 GAAATGAAATGGAATGAACCTGTATCCAGTGTGGAAGCGGGAGAGCGGAGGTGAAG 253  
270 AATCTCTGGAAGGGAGCGGTGTCAGCGGCTCTGACCAATGACTCACCAGCCCTCGTG 329  
254 GACTCTCTGGAAGGGAGCGGTGTCAGCGAGTCTGACCAATGACTCACCAGGCTCTGGTG 313  
330 GGCTCAAAATATAACATTTGGGTGAACCTGTATATCCCTAGATGCCAAAAGGAAGATGCC 389  
314 GGTTCCAATATCACTTTTGGTGAACCTGTGTGTTCCAGATGCCAGAGGAAGATGCT 373  
390 AATGGCAACATAGTCTATGGAAGAACTGCAAGAAATGAGGCTGTTTATCTGATGCCA 449  
374 AATGGCAATATCGTCTATGGAAGAACTGCAAGAAATGATTGGGACTGACCTCTGACCTG 433  
450 TATGTTTACAATCGGACAGCATGTCAGAGACAGTACGCGGGAATGGACCGCCAA 509  
434 CATGCTACAACTGCACTGCAAGGGCGAGATGATGGTGAAGTGGCAGCAGCGGA 493  
510 AGCCATCATACGCTCTTCCCTGATGGAAGAACTTCTCCTACACCCCGGATGGAAGA 569  
494 AGCCAGCATCTCAGGTTCCCGGACAGAGGGCCCTTCCCTCGCCCATGATGGAAGAA 553  
570 TGGAAATTCATCTACGCTTTCCACACACTTGGTCACTATTTCCAGAAATCGGACGATGT 629  
554 TGGAGCTTTGCTACGCTTTTCCACACTTGGCCAGTATTTCCAAAACCTGGTGGTGT 613  
630 TCAGTGAGAGTTCTGTGGAACACAGCCAAATGTGACATTTGGGCTCACTCATGGAAGT 689  
614 TCAGCAGGGTTCCTATAAACACAGTCAACTTGAAGCTGCGGCTCAGGTATGGAAGT 673  
690 ACTGCTACAGAACATGAGCGGCGATATGTTCCATCGCACAAATGGAAGATGTGTAC 749  
674 ACTGCTTCAGAGATCGCGCGGCGCATACATCCCATCTCGAAGTGAAGATGTGTAT 733  
750 GTGGTAACAGATCAGATTCCTGTTGTGATGATGATGTTCCAGAGAAAGATCGAAATCA 809  
734 GTGATAACAGATCAGATCCCTGTATTCGTGACCATGTGCCAGAGAAATGACAGGAATG 793  
810 TCCGACGAACTTCC-CAAGAGTCTCCCATATGTTGATGTCCTGATTCATGATCCT 868  
794 TCTGATGAGATCTTCCCTCAGAGACTCCCTCATGATGCTTCGATGCTCCTCATGATCCC 853  
869 AGCCACTTCTCAATATTTACCAATTAACATCAAGTGGAGCTTCGGGGAATAACTGGC 928  
854 AGCCACTTCTCAAGCATCTCGCATTTCTTCAAGTGGAACTTTGGGGAACAACTGGC 913  
929 CTGTTTGTTCACCAATCATATCTGTGAATCAACGATGATGCTCAATGGAACCTTCAGC 988  
914 CTGTTTGTTCACCAATCATACACTTTGAATCAGATTTATGCTCAATGGAACCTTCAC 973  
989 CTTAACTCTCATGTGAAGCTGCAGCACCAGGACCTGTGCGCCACCGC----- 1037  
974 CTTAACTCTCAGCTGCAAACTGCACTGCGCGGCAATGCCCTTCCCTTCGCTTCGACT 1033  
1038 -----CACCACCCAGACCTTCAAAA 1060  
1034 CGGCTCCACCTTCAACTCGGCGCTCACCCTCGGCTCACCCTGCGCAATTAACA 1093  
1061 CCCACCCCTTCTTAGGACCTGCTGTTGACACCCCTGAGCTGATGAGGATTTCTGAT 1120  
1094 CCTAGCCCTCTTTAATGCCTACTGTTTACAAATCCATGAGCTGATGACATTTCCAAAT 1153

1121 GAAACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATTCACAATTGTAGAGGA 1180  
1154 GAAACTGCCAGATAAACAGATATGGCTACTTTCAGAGCCACCATTCACAATTGTAGAGGG 1213  
1181 ATCTTAGAGTTTAAACATCATCCAGATGACAGAGTCTCTGATGCGGTGCGCATGGCCCTGAA 1240  
1214 ATCTGGAAGTCCAGCATCATGAGATGTCCTCCCATGTCCTCCACACCGCAGCTGCC 1273  
1241 AGTCCCTATAGACTTTGTCGTGACCTGCCAAGGAGATTCCTCCAGGAGTCTGTATCC 1300  
1274 AATCCCTGATGAGACTTCACTGTGACCTGCAAGGGGCCACCCCATGAGAGCTGTACG 1333  
1301 ATCAATTTCTGACCCCACTCGAGATCACCCAGAACACAGTCTCTGACCCCTGTGGATGTG 1360  
1334 ATCATCTCCACCCCACTCGAGATCGCCAGAACCGGTCTGTGACCCCTGTGGCTGTG 1393  
1361 GATGAGATGTCTGTGATCTGTGAGACGAACTTCAATGGTCTGGAGCTGACTGTG 1420  
1394 GATGGGCTGTGCTGTCTGTGAGAAGAGCTTCAATGGTCTGGCACTACTGTGTG 1453  
1421 AACCTACCCCTGGGGATGACACAGCTCGCTCTCACGAGACCTGATTTCTGTTCCT 1480  
1454 AATTTCACTCTGGAGATGATGCAAGCTGTGCTCTCACGAGACCTGATCTCTATCCCT 1513  
1481 GACAGACCCAGCTCGCTTTAAGGATGGCAACAGTGGCCCTGATCTCCGTTGGCTGC 1540  
1514 GGCAGAGCCAGACTCCCTCTGAGAGCAGTGAATGGTGTCTCTGATCTCCATTGGCTGC 1573  
1541 TTGCCATATTTGTCATCTGATCTCTCTGTTGTACAAAAACACAGGAATACAAAC 1600  
1574 CTGGCTGTGCTGTGTCACATGTTTACCATCTTGTGTACAAAAACACAGGCGGTACAG 1633  
1601 CCATAGAAATAGTCTCTGGGAATGTGTGAGAGCAAGGCTGAGTCTTCTCTCAAC 1660  
1634 CCAATAGAAACTCTCCAGGAAACACGCTGAAGGCAAGGCTGAGTCTCTCTCAGC 1693  
1661 CGTGCAAAAGCGGTGTTCTTCCCGGAAACACAGGAAGGATCCGCTACTCAAAACCAA 1720  
1694 CACGGAAGCGCCCTTCTCCGAGGAGACACAGGAGAGGATCCATGCTCCAGGACAAG 1753  
1721 GAATTTAAGGAGTTCTTAAATTTGACCTTGTGTTCTGAGAGCTCACTTTTCAGTCCAT 1780  
1754 CCAAGGACACTCTAAGTCTTTGGCTTCCCTCTGACAGGAACCCACTCTTCTGTG-CAT 1812  
1781 TGATGTGAGATGTGCTGA-----GTGGCTATTAACCTTTTCTTAAAGATTAATGTTA 1836  
1813 GTATGTGAGCTGTGAGAGTATGTGGCTGGAACTGTGTTCTCTAAGGATTAATGTTAA 1872  
1837 AATAGATATTTGTTGGGGAAGTTGAATTTTATAGGTTAAA 1881  
1873 AATGATATCTGGCTTAGGAGTGTGGTTAAATAGCATTTTATA 1917

## RESULT 14

BC026375  
LOCUS BC026375  
DEFINITION Mus musculus glycoprotein (transmembrane) mmb, mRNA linear ROD 06-OCT-2003  
MGC:31121 IMAGE:4164706, complete cds.  
ACCESSION BC026375  
VERSION BC026375.1 GI:20071338  
KEYWORDS MGC.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 2299)  
Strausberg,R.L., Feingold,B.A., Grouse,L.H., Derge,J.G.,  
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,  
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,  
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,  
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,  
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,

Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Lequellano, N.A., Peters, G.J., Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalilus, D.E., Schnerker, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

23388257

PUBMED

12477932

2 (bases 1 to 2299)

Strausberg, R.

Direct Submission

Submitted (02-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) [mcdpaxil.stanford.edu](mailto:mcdpaxil.stanford.edu)

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAC Plate: 43 Row: d Column: 20

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

#### FEATURES

source

Location/Qualifiers

1..2299

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="MGC:31121 IMAGE:4164706"

/tissue\_type="Salivary gland, 10 week old female mouse"

/clone\_lib="NCI CGAP\_SG2"

/lab\_host="DH10B"

/notes="vector: pCMV-SPORT6"

1..2299

/gene="Gpmb"

/db\_xref="LocusID:93695"

/db\_xref="MGI:1934765"

80..1804

/codon\_start=1

/product="glycoprotein (transmembrane) nmb"

/protein\_id="AAH26375.1"

/db\_xref="GI:20071339"

/db\_xref="LocusID:93695"

/translation="MESLGVGLFLLAAGLPQAKRFRDVLGHQYFNHMRHNL  
RWSDENDEHLVPMRGDGRWDSWEGRVQAVLTSDSPALVGSNTFVNLV  
PRCKEDANGNIYVENCNDLGLTSLHYNWNTAGDDWEDGTSRSHLRFDR  
PFRPHGKWSFVYVHTLGOYFQKLGRCARSVINTVNLTAGPQMEVTVRYGR  
AVIPISKVDVVLTDIPIEVFTMSQKDRNLSDELFLDLPIEDVLHDPHFLND  
SALSVKWNFGDNTGLFVSNHNLNHTYVLNGTFNLNLVTQTAVPGPCPPSPPPP  
STPPSPSPPLTSP  
EVSIMQIDVPMPTQPMANLMDFTVCKGATPMEACTIISDPTQIAQNRVCSFVAV  
DGLCLLSVRRAFNGSGTCVNFNTLGLDASLALTSLISIPGKDPSPPLRVNGVLISI

#### gene

#### CDS

GCLAVLTWVITLLYKHKAYKPIGNCPRNTVKGKLSVLLSHAKAPFRGQEKDPL  
LQDKRPTL"

854..1039

/note="PKD; Region: Repeats in polycystic kidney disease 1  
(PKD1) and other proteins"

/db\_xref="CDD:smart0089"

ORIGIN

Query Match 39.5%; Score 1049.8; DB 10; Length 2299;  
Best Local Similarity 74.0%; Pred. No. 8.4e-248;  
Matches 1409; Conservative 0; Mismatches 442; Indels 54; Gaps 4;

QY 30 CTTGGTGCCTGGCGTCCGTGAGAAATTCAGCATGGAATCTCTCTACTATTTCTTCTGGGATTT 89  
DB 50 CTTGACTGGTTGCAGGCGCTCGGAGTCAGCATGGAAGTCTCTCGGGGTCTCGGATTT 109

QY 90 CTGCTCCCTGGCTGCAAGATTGCGCACTTGATGCCGCAAAACGATTTCATGATGTGCTGGGC 149  
DB 110 CTGCTGGCTGGCTGAGGACTGCTCTCCAGCTGCCAGCGATTTCGTGATGTGCTGGGC 169

QY 150 AATGAAGACCTTCTGCTTACATGAGGAGCACAAATCAATTAATTAATGGCTGCTCTCTGAT 209  
DB 170 CATGAACAGTATCCCAATCATTGAGAGGACCAACCAATACGTGGCTGCTCTCGGAT 229

QY 210 GAAATGACTGGAATGAAATACTCTACCACTGTGGAAGCGGGAGACATCAGGTGAAA 269  
DB 230 GAAATGAATGGGATGAACACCTGTATCCAGTGTGGAGGAGGAGCGGAGGTGGAG 289

QY 270 AACTCTCTGGAGGAGCGCGTGTGAGCGGTCTGACAGTGTCTACAGTGTCTACAGCGCTCGTG 329  
DB 290 GACTCTCTGGAGGAGCGCGTGTGAGCGGTCTGACAGTGTCTACAGTGTCTACAGCGCTCTGTG 349

QY 330 GGCTCAATATAACATTTGGCGTGAACCTGATATTCCTAGATGCCAAGAGGAGATGCC 389  
DB 350 GGTCCCAATATCACCTTTGTTGGTGAACCTGTTCCCGATGCCAAGAGGAGATGCT 409

QY 390 AATGGCAACATAGTCTATGAGAGAACTGCAGAAATCAGGCTGGTTTATCTGCTGATCCA 449  
DB 410 AATGGCAATATCGTCTATGAGAGAACTGCAGGAATGATTTGGGACTGACTCTGACCTG 469

QY 450 TATGTTTACAACTGGGACAGCATGTCAGAGGACAGTGCAGGGGAAATGCAACCGGCAA 509  
DB 470 CATGCTTACAACTGGGACAGCATGTCAGGGGAGATGATGTTGTTGGAAGATGTCACAGCGA 529

QY 510 AGCCATCAATAACGCTCTCCCTGATGGAAACCTTTTCCACACCCCGATCGAGAGAGA 569  
DB 530 AGCCAGCATCTCAGTTTCCCGACAGAGGCGCTTCCCTCGCCCCCATGATGGAAGAA 589

QY 570 TGAATTTCACTACCTCTTCCACACACTTGGTTCAGTATTTCCAGAAATTTGGGACGATGT 629  
DB 590 TGGAGCTTTGTCTACGCTTTTCCACACACTTGGCCAGTATTTCCAAATCTGGTTCGGTGT 649

QY 630 TCAGTGAGAGTTTCTGTGAACACAGCAATGTGACATTTGGGCTCAACTCATGGAAGTG 689  
DB 650 TACGACGGGTTTCTATAAACACAGTCAACTTGACAGTGGGCTCGGCTCATGGAAGTG 709

QY 690 ACTGTCTACAGAGACATCGACGGGCATATGTTCCCATCGACAGTGAAGATGTGTAC 749  
DB 710 ACTGTCTTGAAGATACGCGCGGGCATACATTCCTCATCTCGAAGGTGAAGATGTGTAT 769

QY 750 GTGGTAAACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAAAGAACCATCGAAATCCA 809  
DB 770 GTGATAACAGATCAGATTCCTGTATGTTGACCATGTCTCCAGAGATACAGGACTTG 829

QY 810 TCCGAGCAACCTTC-CAAAGATCTCCCATATGTTTGTGATGTTGATGTTGATGATCT 868  
DB 830 TCTGATGAGATCTTCTCAGAGACCTCCCACTCGTCTCGATGCTCTCTATTCTGATGCC 889

QY 869 AGCCACTTCTCAATTTATTTCTACCACTTAACATGAGTGGAGCTTGGGATTAATGATGCC 928  
DB 890 AGCCACTTCTCAAGACTCTGCCATTTCTCAAGTGGAACTTTGGGACAACTGCG 949

QY 929 CTGTTGTTTCCCAACCATCATACTGTGAATCAACAGTATGCTCAATGGAACCTTCAGC 988

950 CTGTTTGTCTCCAAATATCACTTTGAAATCACTTATGTGCTCATGAACTTCAAC 1009  
989 CTTAACTCTACTGTGAAGCTGCAGACCCAGGACCTTTGTCCGCCACCGC----- 1037  
1010 CTTAACTCTACCGTGAAGCTGCAGTGCCTCCGCGCATGCGCTCCCTTGGCTTCGACT 1069  
1038 -----CACCACACCCAGACCTTCAAA 1060  
1070 CCGCTCCACCTTCAACTCCGCCCTCACTCCGCCCTCACCTCGCCACATATCAACA 1129  
1061 CCCACCCCTCTTTAGGACCTGTGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGAT 1120  
1130 CCTAGCCCTCTTTAATGCCCTACTGTGTTAATAATCCATGGAGCTGAGTGACATTTCCAAT 1189  
1121 GAAACTGCCAGATTAACAGATATGGCACTTTCAAGCCACCATCAACAATTTGAGGGA 1180  
1190 GAAACTGCCAGATTAACAGATATGGCTACTTCAGAGCCACCATCAACAATTTGAGGGA 1249  
1181 ATCTTAGAGGTAAATCATCATCCAGATGACAGACCTTCCTGATGCGCGTCCCATGGCTGAA 1240  
1250 ATCTTGAAGTACGATCATGATAGATGATGCTCCCATGCTCCCATGAGGCTGCTGCC 1309  
1241 AGCTCCCTTAATAGACTTTGTGCTGACCTGCCAAGGGAGCATTTCCACGAGGCTCTGTACC 1300  
1310 AACTCCCTGTAGGACTTCACTGTGACCTGCAAGGGGCCACCCCATGGAAGCTGTAGC 1369  
1301 ATCAATTTCTGACCCCTGCGAGATCAACCCAGAAACAGTGTGAGCCCTGTGTGATGTG 1360  
1370 ATCAATCTCCGACCCCTGCGAGATCGCCAGAACCCGGGTCTGAGGCTGTGCTGTG 1429  
1361 GATGAGATGTCTGCTGACTGTGAGACGAACTTCAATGGTCTGGGACGTACTGTGTG 1420  
1430 GATGGGCTGTGCTGCTGTGTGAGAGAGCTTCAATGGTCTGTGGACCTTACTTGTGTG 1489  
1421 AACCTCACCTGGGGGATGACAAAGCTGGCTCTCACAGACACCTGATTTCTGTCTCT 1480  
1490 AATTTCACTCTGGAGATGATGCAAGCTGGCCCTCACAGACACCTGATCTATCCCT 1549  
1481 GACAGAGACCCAGCTCGCTTTAAGATGGCAACAGTGGCCCTGATCTCCGTTGGCTGC 1540  
1550 GCAAAAGACCCAGACTCCCTCTGAGAGCAGTGAATGGTGTCTGATCTCCATGGCTGC 1609  
1541 TTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAAACACAAAGAAATACAA 1600  
1610 CTGGCTGTCTGTCAACATGTTTACATCTTGTCTGTACAAAACACAAAGGCTACAA 1669  
1601 CCAATAGAAAATAGTCTCTGGGAATGTGGTCAGAAACAAAGGCTGAGTGTCTTTCTAAC 1660  
1670 CCAATAGGAACTGCCCCAGGACACAGGTCAAGGGCAAGGCTGAGTGTCTCTCTCAGC 1729  
1661 CGTGAAAAGCCGTGTCTTCCGGGAAACACAGGAAAAGATCCGCTACTCAAAACCAA 1720  
1730 CACGCGAAAGCCCGTCTTCTCCGAGGAGACACAGGAAAGATCCATTGCTCCAGGACAAG 1789  
1721 GAAATTAAGGAGTTTCTTAAATTTGACCTTGTCTGAAGCTCACTTTTCAAGTCCCAT 1780  
1790 CCAAGGACACTTAAGTCTTTGGCCCTTCCCTCTGACGAGAACCCACTTCTCTGTG-CAT 1848  
1781 TGATGTGAGATGTCTGGA-----GTGGCTATTAACTTTTTCCTTAAAGATTTATTGTA 1836  
1849 GTATGTGAGCTGTGCAGAGATGATGGCTGGGAACCTGTGTTCTCTAAGGATTTATTGTA 1908  
1837 AATAGATATTGTGTTTGGGGAAGTTGAATTTTATTAGTTTAA 1881  
1909 AATGTATATCGTGGCTTAGGGAGTGTGGTTAAATAGCATTTTGA 1953

RESULT 15  
AR156839  
LOCUS  
DEFINITION Sequence 27 from patent US 6242419.  
ACCESSION AR156839

AR156839.1 GI:15125543  
Unknown.  
ORGANISM  
Unclassified.  
REFERENCE  
1 (bases 1 to 2213)  
Strachan,L., Sleeman,M., Abernethy,N., Onrust,R., Kumble,A. and  
Murison,G.  
TITLE  
Compositions isolated from stromal cells and methods for their use  
JOURNAL  
Patent: US 6242419-A 27 05-JUN-2001;  
FEATURES  
Location/Qualifiers  
1..2213  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Query Match 39.4%; Score 1047.8; DB 6; Length 2213;  
Best Local Similarity 74.1%; Pred. No. 2.6e-247;  
Matches 1404; Conservative 0; Mismatches 437; Indels 54; Gaps 4;  
QY 40 TGCCTCCGTGAGAATTTCAGCATGGAATGTCTCTACTATTTCTCTGGGATTTCTGCTCTGG 99  
Db 3 TGCAGGCGCTCGGAGTCAGCATGGAAGTCTCTCGGGGTCTCTGGGATTTCTGCTCTGG 62  
QY 100 CTGCAAGATTCCTACTGATGCCGCCAAACGATTTTCATGATGTCTGGGCAATGAAGAC 159  
Db 63 CTGCAGGACTGCCTCTCCAGGCTGCCAAGCGATTTTCGTGATGTCTGGGCCATGAACAGT 122  
QY 160 CTTCTGCTTACATGAGGAGCACAAATCAATTAATTAATGGCTGTCTCTCATGAAAATGACT 219  
Db 123 ATCCCAATTCATGAGAGAGCACAAACDANTTACGTGGGTGTCTTCGGATGAAATGAAT 182  
QY 220 GGAATGAAAACTCTTACCAGTGTGGAGCGGGAGACATGAGTGTGAAAAAACTCCTCGA 279  
Db 183 GGGATGAACACTGTATCCAGTGTGGAGAGGGAGACGGCAGGTGGAAGGACTCCTCGG 242  
QY 280 AGGAGGCGGTGTGAGGGGTCTGACAGTGTGACAGTGTCTCCAGCCCTCGTGGCTCAATA 339  
Db 243 AAGGAGGCGGTGTGAGGCGAGTCTCTGACAGTGTCTCCAGGCTCTGGTGGGTTCGAATA 302  
QY 340 TAAACATTTGCGGTGAACCTGATATTCCTAGATGCCAAAGGAAGATGCCAATGSCAA 399  
Db 303 TCACCTTTGTGGTGAACCTGTGTTCCTCAGATGCCAGATGCCAGAAAGATGCTAATGCAATA 362  
QY 400 TAGTCTATGAGAAGAACTGCAGAAATGAGGCTGTGTTTATCTGCTCATCATATGTTTACA 459  
Db 363 TCGTCTATGAGAAGAACTGCAGAAATGATTTTGGGACTGACCTCTGACCTGCTATCTACA 422  
QY 460 ACTGACAGCATGTGTCAGAGACAGTGCAGGGGAAAATGGCCCGGCAAGCCATCATATA 519  
Db 423 ACTGGAAGTGCAGGGGAGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 482  
QY 520 AGTCTTCCCTCATGSGAAACCTTTTCTCTCACCACCCCGGATGGAGAAAGATGGAATTTCA 579  
Db 483 TCAGGTTCCCGACAGGAGGCTCTCCCTCGCCCTCATGATGGAAGAAATGGAGCTTTG 542  
QY 580 TCTACGCTTTCACACACTTGTGTGATATTTCCAGAAATTTGGAAGATGTTGATGAGTGA 639  
Db 543 TCTACGCTTTCACACACTTGTGTGATATTTCCAAAGATTTGTTGTTGTTGTTGTTGTTGTT 602  
QY 640 TTTCTGTGAACACAGCAATGTGACACTTTGGGCTCAACTCATGGAAGTGAAGTGTCTACA 699  
Db 603 TTTCTATAAACACAGTCAACTTTGACAGTGGGCTCTCAGGTCATGGAAGTGAAGTGTCTTTC 662  
QY 700 GAAGACATGAGCGGCATATGTTCCCATCGCAAGTGAAGATGTTGATGTTGATGTTGTTGTT 759  
Db 663 GAAGATACGCGCGGATACATTTCCCATCTCGAAGGTGAAGATGTTGATGTTGATGTTGTT 722  
QY 760 ATCAGATTTCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 819  
Db 723 ATCAGATTTCTGTATTTCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 782  
QY 820 CCTTCC-CAAGATCTCCCCATTTATGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 878

Db 1862 GTGGCTTAGGGAGTGTGGTTAAATAGCATTTTAGA 1899

Search completed: September 10, 2004, 07:08:45

Job time : 16437 secs

783 TCTTCTCAGAGACCTCCCATCGTCTTCGATGTCTCATTCATGATCCAGCACTCC 842

879 TCAATTATTTACCATTAACATAAGTGGAGCTTCGGGGATATATCTGGCCTCTTTGTTT 938

843 TCAAGGACTCTGCCATTTCTCAAGTGGAACTTTGGGGACAACACTGGCCTGTTTGTCT 902

939 CCACCAATCATACTGTGAATCACACGCTATGTCTCAATGGAACTTTAGCCTTTAACCTCA 998

903 CCAACAATCACACTTTGNAATCACACTTATGTCTCAATGGAACCTTCAACCTTTAACCTCA 962

999 CTGTGAAGCTGCAGCAACAGGACCTTG----- 1026

963 CCGTGCAAACTGAGTGGCCGGGCCATGCCCCCTCCCTTCGCTTCGACTCCGCTCCAC 1022

1027 -----TCCGCCACCGCCACACCCAGCACTTCAAAACCCACCCCTT 1070

1023 CTTCAACTCGCCCTACACTCCGCCCTCACCTCTGCCCACTTATCAACACTAGCCCT 1082

1071 CTTTAGGACCTGCTGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCC 1130

1083 CTTTAAATGCTACTGTGTTACAAATCCATGGAGCTGAGTGACATTTCCATGAAACTGCC 1142

1131 AGATTACAGATATGGCCACTTTCAGCCACCACTACAAATTGTAGAGGGAATCTTAGAGG 1190

1143 GAATAAACAGATATGGCTACTTCAGAGCCACCACTACAAATTGTAGAGGGGATCCTGGAAG 1202

1191 TTAACATATCCAGATGACAGAGCTCTGATGCGGTGCCATGGCTTGAAGCTCCCTAA 1250

1203 TCAGCATATGCAATAGAGATGTCCCAATGCCACACCGAGCCTGCCAATCCCTGA 1262

1251 TAGACTTTGTGTCGACCTGCCAAGGGAGCAATCCACGGAGGTCGTGTACCATCAATTTCTG 1310

1263 TGGACTTCACTGTGACCTGCAAGGGGCCACCCCTCATGGAAGCCTGTACCATCATCTCCG 1322

1311 ACCCACTCGGAGATCAACCCAGAAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGT 1370

1323 ACCCACTGCGAGATCGCCAGAACCGGGTCTGCAGCCCTGTGGCTGTGGATGGGCTGT 1382

1371 GTCTGCTGACTGTGAGAGCACTTCAATGGTCTGGAGCTACTGTGTGAACCTCACCC 1430

1383 GCCTGCTGTCTGTGAGAAGAGCCTTCAATGGTCTGGCACTACTGTGTGAATTTCACTC 1442

1431 TGGGGATGACACAAGCCTGGCTCTCACGAGCACCTGATTTCTGTTCCTGCAGAGAGCC 1490

1443 TGGGAGATGATCAAGCCTGSCCTCACAGCACCTGATCTCTATCCCTGGCAAGACC 1502

1491 CAGCCTCGCCTTTAAGGATGGCAAAACAGTGCCTGATCTCCGTTGGCTGTGGCCATAT 1550

1503 CAGACTCCCTCTGAGAGCAGTGAATGGTGTCTCTGATCTCCATTTGGCTGGCTGGCTGTGC 1562

1551 TTGTCACCTGTGATCTCCCTCTTGGTGACAAAACACAAAGGAATCAACCCCAATAGAAA 1610

1563 TTGTCAACATGGTATACCATCTGCTGTACAAAACACAAAGGGTCAAGCCCAATAGGAA 1622

1611 ATAGTCTGGGAATGTGGTCAAGAGCAAGGCTGAGTGTCTTCTCAACCGTGCAAAAG 1670

1623 ACTGCCCCAGGAACACCGTCAAGGGCAAGGCTGAGTGTCTCTCTCAGCCACGCGAAG 1682

1671 CCGTGTCTTCCCGGGAAACAGGAAAGGATCGGCTACTCAAAAACCAAGAAATTTAAG 1730

1683 CCGCGTCTTCCGAGGAGACCGAGGAAGGATCCATTTGCTCCAGGCAAGCCAGGACAC 1742

1731 GAGTTTCTTAATTTCCAGCTTGTCTGAGCTCACTTTTCAGTGCCATTGATGTGAGA 1790

1743 TCTAAGTCTTTGGCCTTCCCTCTGACAGGAACCACTCTTCTGTG-CATGTATGTGAGC 1801

1791 TGTGCTGA-----GTGGCTATTAACTTTTCTCTAAAGATTATGTTAAATAGATATT 1846

1802 TGTGCAGAGTATGTGGCTGGAACTGTGTCTCTAAGGATTATGTTAAATGTATATC 1861

1847 GTGTTTGGGAAGTGAATTTTATAGGTTAAA 1881







A:Title: Identification of a human melanoma antigen recognized by tumor-infiltrating lymphocytes  
A:Reference number: A55753; MUID:94294401; PMID:8022805  
A:Accession: A55753  
A>Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-161, 'P', 163-592, 594-662 <RAW>  
C:Keywords: glycoprotein

Alignment Scores:  
Pred. No.: 6,05e-43 Length: 662  
Score: 593.50 Matches: 167  
Percent Similarity: 39.45% Conservative: 93  
Best Local Similarity: 25.34% Mismatches: 200  
Query Match: 12.24% Indels: 200  
DB: 2 Gaps: 17

US-10-039-272-1 (1-2661) x 138400 (1-662)

QY	66	TGTCTCTACTATTTCCTGGATTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCGGC	125
DB	8	CysLeu-----LeuHisLeuAlaValIleGlyAlaLeuLeuAlaValGlyAlaThr	24
QY	126	AAA-----CGATTTCATGATGCTGGCAATGAAAGACCTTCTCTTACATGAGGAG	179
DB	25	LysValProArgAsnGlnAspTrpLeuGlyValSerArg-----	37
QY	180	CACAATCAATTAATGGCTGCTCTCTGATGAAATGACATGGAATGAAATCACTTACCCA	239
DB	38	-----GlnLeuArg-----ThrLysAlaTrpAsnArgGlnLeuTrpPro	50
QY	240	GTGTGAAGCGGGAGACATGAGTGGAAAAAATCCTCTGGAAGGGAGCGCTGTGACGGCG	299
DB	51	GluTrpThr-----GluAlaGlnArgLeuAspCysTrpArgGlyGlyGlnValSerLeu	68
QY	300	GTCTGACGAGTCACTACAGCCCTGTGGCTCAAAATATAACATTTGCGGTGAACCTG	359
DB	69	LysValSerAsnAspGlyProThrLeuIleGlyAlaAsnAlaSerPheSerIleAlaLeu	88
QY	360	ATATTCCTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTTATGAGAAGAACTGC	419
DB	89	AsnPheProGlySerGlnLysValLeuProAspGlyGlnValIleTrpValAsnAsnThr	108
QY	420	AGAAATCAGGCTGTTTATCTGCTGATCCATATGTTTACAAC-----TGG---	464
DB	109	-----IleLeuAsnGlySerGlnValTrpGly	117
QY	465	-----ACAGCATGTCAGAGGACAGTACAGGGGAAATGGCACCGCCAAAGCCATCAT	518
DB	118	GlyGlnProValTrpProGlnGluThrAspAla-----	129
QY	519	AACGCTCTCCCTGATGGAAACCTTTCTCACCACCGGATGGAGAGATGGAATTC	578
DB	130	CysIlePheProAspGlyGlyProCysProSerGlySerTrpSerGlnLysArgSerPhe	149
QY	579	ATCTAGCTCTTCACACACTTGTGCTAGTATTTCCAGAAATTTGGACGATGTTCCAGTGA	638
DB	150	ValTrpValTrpLysThrTrpGlyGlnTrpGlnValLeuGlyGlyProValSerGly	169
QY	639	GTTCCTGTGAACACAGCAATGTGACATTTGGCCCTCACTCATGGAAGTACTGTCTAC	698
DB	170	LeuSerIleGlyThrGlyArgAlaMetLeuGlyThrHisThrMetGluValThrValTrp	189
QY	699	AGAAGACATGGA---CGGGCATATGTTCCCATCGCAACAGTGAAGATGTGACGTGGA	755
DB	190	HisArgArgGlySerArgSerTrpValProLeuAlaHisSerSerSerAlaPheThrIle	209
QY	756	ACAGATCAGATTCCTGTGTTTGTGACTATGTTCCAGAAAGAACGATCGAAATTCATCCGAC	815
DB	210	ThrAspGlnValProPheSerValSerValSerGlnLeuArgAlaLeuAspGlyGlyAsn	229
QY	816	GAACACTTCC---AAAGATCTCCCATATGTTTGTGCTGATTCATGATCTTAGCCAC	874
DB	230	LysHisPheLeuArgAsnGlnProLeuThrPheAlaLeuGlnLeuHisAspProSerGly	249
QY	875	TTCTCAATTATTCTTACCATTAACTACAGTGGAGCTTCGGGATAATAATGCGCTGTTT	934
DB	250	TyrLeuAlaGluAlaAspLeuSerTrpTrpAspPheGlyAspSerSerGlyThrLeu	269
QY	935	GTTCACCAATCATCTGTGAATCACAGTATGTCTCAATGGAACCTTCAGCCTTAAC	994
DB	270	IleSerArgAlaLeuValValThrHisThrTrpLeuGluProGlyProValThrAlaGln	289
QY	995	CTCACTGTGAAGTGCAGCACCA-----GGACCTTGTTCGGCCACCGCCA---	1039
DB	290	ValValLeuGlnAlaIleProLeuThrSerCysGlySerProValProGlyThr	309
QY	1039	-----	1039
DB	310	ThrAspGlyHisArgProThrAlaGluAlaProAsnThrThrAlaGlyGlnValProThr	329
QY	1039	-----	1039
DB	330	ThrGluValValGlyThrThrProGlyGlnAlaProThrAlaGluProSerGlyThrThr	349
QY	1039	-----	1039
DB	350	SerValGlnValProThrThrGluValIleSerThrAlaProValGlnMetProThrAla	369
QY	1039	-----	1039
DB	370	GluSerThrGlyMetThrProGluLysValProValSerGluValMetGlyThrThrLeu	389
QY	1039	-----	1039
DB	390	AlaGluMetSerThrProGluAlaThrGlyMetThrProAlaGluValSerIleValVal	409
QY	1039	-----	1039
DB	410	LeuSerGlyThrThrAlaAlaGlnValThrThrGluTrpValGluThrThrAlaArg	429
QY	1040	-----CCACCACCCAGACCTTCAAAACCC-----	1063
DB	430	GluLeuProIleProGluProGluGlyProAspAlaSerSerIleMetSerThrGluSer	449
QY	1064	-----ACCCCTTTTAGGACCTGCTGGTGACACCCCTGAGCTGATGAGTATCTCGAT	1120
DB	450	IleThrGlySerLeuGlyProLeuLeuAspGlyThrAlaThrLeuArgLeuValLysArg	469
QY	1121	GAA-----AACTGCCAGATTACAGATATGGCCACTTCAAGCCACCATCACAAAT	1171
DB	470	GlnValProLeuAspCysValLeuTrpArgTrpGlySerPheSerValThrLeuAspIle	489
QY	1172	GTAGAGGGAATCTTAGAGGTTAATCATCCAGATGACAGACGCTCTGATGCCGGTGCCA	1231
DB	490	ValGlnGly-----IleGluSerAlaGluIleLeuGlnAlaValPro	503
QY	1232	TGGCCTGAAAGCTCCCTAATAGACTTTGCTGACCTGCAAGGGAGCATTCACCGGAG	1291
DB	504	SerGlyGluGlyAspAlaPheGluLeuThrValSerCysGlnGlyGlyLeuProLysGlu	523
QY	1292	GTCTGTACCATCATTTCTGACCCCTCGAGATCACCAGAACACAGTCTCAGCCCT	1351
DB	524	AlaCysMetGluIleSerSerProGlyCysGlnProProAlaGlnArgLeuCysGlnPro	543
QY	1352	GTGAGTGTGGATGAGATGTCTGTGCTGACTGTGAGACGAACTTC-----AATGGCTGGG	1408
DB	544	ValLeuProSerProAlaCysGlnLeuValLeuHisGlnIleLeuLysGlySerGly	563
QY	1409	ACGTACTGTGTGAACCTCACTCCCTGGGGGATGACAAAGCTGTGCTCTCAGACACCTG	1468
DB	564	ThrTrpCysLeuAsnValSerLeuAlaAspThrAsnSerLeuAlaValSerThrGln	583
QY	1469	ATTTCGTCTCTGACAGACACCCAGCTCGCCTTTAAGGATGCAAAACAGTGCCTGATC	1528
DB	584	LeuIleMetProGlyGlnGluAlaGlyGly-----LeuGlyGlnValProLeuIle	600

QY 1529 TCGTGGTGGTGGCCATATTTGTCACGTGTGATCTCCCTCTGGTGTACAAAAA 1585  
Db 601 ValGlyIleLeuLeuValLeuMetAlaValValLeuAlaSerLeuIleValArgArg 619  
RESULT 4  
A41234  
melanocyte-specific protein Pmel-17 precursor - human  
C:Species: Homo sapiens (man)  
C>Date: 19-Jun-1992 #sequence\_revision 19-Jun-1992 #text\_change 30-Sep-1993  
C/Accession: A41234  
R/Kwon, B.S.; Chittamaneni, C.; Kozak, C.A.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.;  
Proc. Natl. Acad. Sci. U.S.A. 88, 9228-9232, 1991  
A/Title: A melanocyte-specific gene, Pmel 17, maps near the silver coat color locus on m  
A/Reference number: A41234; MUID:92021023; PMID:1924386  
A/Accession: A41234  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-668 <KMO>  
A/Cross-references: GB:M77348  
Alignment Scores:  
Pred. No.: 1-65e-42 Length: 668  
Score: 588.50 Matches: 170  
Percent Similarity: 39.61% Conservative: 93  
Best Local Similarity: 25.60% Mismatches: 198  
Query Match: 12.13% Indels: 204  
DB: 2 Gaps: 19  
US-10-039-272-1 (1-2661) x A41234 (1-668)  
QY 66 TGCTCTPACTATTTCCTGGGATTTCTGCTGCTGCAAGATTGCCACTTGATCGCGCC 125  
Db 8 CysLeu-----LeuHisLeuAlaValIleGlyAlaLeuLeuAlaValGlyAlaThr 24  
QY 126 AAA-----CGATTTCATGATGCTGGGCAATGAAGACCTCTGCTTACATGAGGAG 179  
Db 25 LysValProArgAsnGlnAspTrpLeuGlyValSerArg----- 37  
QY 180 CACAATCAATTAATGCTGCTCTCTGATGAAATGACCTGGAATGAAACTTACCCA 239  
Db 38 -----GlnLeuArg-----ThrLysAlaTrpAsnArgGlnLeuTyPro 50  
QY 240 GTGTGAAGCGGGGACACATGAGTGGGAAAACTCTCGAAGGAGCGGTGGCAGCG 299  
Db 51 GluTrpThr-----GluAlaGlnArgLeuAspCysTrpArgGlyGlnValSerLeu 68  
QY 300 GTCTCACACAGTCACTACAGCCCTCGTGGGTCAAATATAACATTGGGTGAACCTG 359  
Db 69 LysValSerAsnAspGlyProThrLeuIleGlyAlaAsnAlaSerPheSerIleAlaLeu 88  
QY 360 ATATTCCTTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAGAAGTGC 419  
Db 89 AsnPheProGlySerGlnLysValLeuProAspGlyGlnValIleTrpValAsnAsnThr 108  
QY 420 AGAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAAC-----TGG--- 464  
Db 109 -----IleIleAsnGlySerGlnValTrpGly 117  
QY 465 -----ACAGCATGTCAGAGGACAGTGAAGGAAAAATGGCAGCGCAAGCCATCAT 518  
Db 118 GlyGlnProValTyProGlnGluThrAspAla----- 129  
QY 519 AAGCTCTCCCTGATGGGAAACCTTTCTCACCACCCCGGATGGAGAAGATGGAAATTC 578  
Db 130 CysIlePheProAspGlyGlyProCysProSerGlySerTrpSerGlnLysArgSerPhe 149  
QY 579 ATCTAGCTCTCCACACCTGCTAGTATTTCAGAAATTCGAGCGAGTTCAGTGAGA 638  
Db 150 ValTyValTrpLysThrTrpGlyGlnTyTrpGlnValLeuGlyGlyProValSerGly 169  
QY 639 GTTTCGTGAACACAGCCCAATGTGACATTGGGCTCAACTGATGAAGTACTGTCTAC 698  
Db 170 LeuSerIleGlyThrGlyArgAlaMetLeuGlyThrHisThrMetGluValThrValTy 189

QY 699 AGAAGACATGCA---CGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGTACGTGTA 755  
Db 190 HisArgArgGlySerArgSerTyValProLeuAlaHisSerSerSerAlaPheThrIle 209  
QY 756 ACAGATCAGATTCCTGTTGTTGTCAGTATGTTCCAGAGAAGACGATCGAAATTCATCCGAC 815  
Db 210 ThrAspGlnValProPheSerValSerValSerGlnLeuArgAlaLeuAspGlyGlyAsn 229  
QY 816 GAAACCTTCCC-AAAGATCTCCCATATGTTGTCCTGCTGATTCATGATCCTAGCCAC 874  
Db 230 LysHisPheLeuArgAsnGlnProLeuThrPheAlaLeuGlnLeuHisAspProSerGly 249  
QY 875 TTCCTCAATTATCTACCAATTAACAAGTGGAGCTTCGGGGATAATATCTGGCTGTTT 934  
Db 250 TyrLeuAlaGluAlaAspLeuSerTyTrpAspPheGlyAspSerSerGlyThrLeu 269  
QY 935 GTTTCACCAATCATACTGTGAATCACAGTATGCTCAATGGAACTTCAGCCTTAAC 994  
Db 270 IleSerArgAlaProValValThrHisThrTyLeuGluProGlyProValThrAlaGln 289  
QY 995 CTCACCTGTGAAGCTGCAGCACCA-----CGACCTTCTCCGCCACCGCCA--- 1039  
Db 290 ValValLeuGlnAlaAlaIleProLeuThrSerCysGlySerProValProGlyThr 309  
QY 1039 ----- 1039  
Db 310 ThrAspGlyHisArgProThrAlaGluAlaProAsnThrThrAlaGlyGlnValProThr 329  
QY 1039 ----- 1039  
Db 330 ThrGluValValGlyThrThrProGlyGlnAlaProThrAlaGluProSerGlyThrThr 349  
QY 1039 ----- 1039  
Db 350 SerValGlnValProThrThrGluValIleSerThrAlaProValGlnMetProThrAla 369  
QY 1039 ----- 1039  
Db 370 GluSerThrGlyMetThrProGluLysValProValSerGluValMetGlyThrThrLeu 389  
QY 1039 ----- 1039  
Db 390 AlaGluMetSerThrProGluAlaThrGlyMetThrProAlaGluValSerIleValVal 409  
QY 1039 ----- 1039  
Db 410 LeuSerGlyThrThrAlaAlaGlnValThrThrThrGluTrpValGluThrThrAlaArg 429  
QY 1040 -----CCACCACCCAGACCTTCAAAACCC----- 1063  
Db 430 GluLeuProIleProGluProGluGlyProAspAlaSerSerIleMetSerThrGluSer 449  
QY 1064 ---ACCCCTCTTAGGACCTGCTGGTGACACCCCTCGAGCTGAGTGGATTCTCGAT 1120  
Db 450 IleThrGlySerLeuGlyProLeuLeuAspGlyThrAlaThrLeuArgLeuValLysArg 469  
QY 1121 GAA-----AACTGGCCAGATTAAACATATCCAGATGACACAGCTCTGATGCGGTGCCA 1171  
Db 470 GlnValProLeuAspCysValLeuTyArgTyGlySerPheSerValThrLeuAspIle 489  
QY 1172 GTAGAGGAATCTTAGAGGTTAACTATCATCCAGATGACACAGCTCTGATGCGGTGCCA 1231  
Db 490 ValGlnGly-----IleGluSerAlaGluIleLeuGlnAlaValPro 503  
QY 1232 TGGCTGAAAGCTCCCTAATAGACTTTCGCGACTGCTCCAGGAGGAGCATTCACGAGG 1291  
Db 504 SerGlyGluGlyAspAlaPheGluLeuThrValSerCysGlnGlyGlyLeuProLysGlu 523  
QY 1292 GTCTGTACCATCATTTCTGACCCCACTCGGAGATCACCAGAACACAGTCTGACGCTT 1351  
Db 524 AlaCysMetGluIleSerSerProGlyCysGlnProAlaGlnArgLeuCysGlnPro 543

1352	GTGGATGTCGATGAGATGTCCTGCTGACTGTGAGACGAACCTTC---	AATGGTCTGGG	1408
QY			
Dd			
544	LeuLeuProSerProAlaCysGlnLeuValLeuHisGlnIleLeuIysGlyIysGlySerGly	563	
QY	ACCTACTGTGTGAACCTCACCTCCCTGGGGTGACACAAAGCTGCCTCTCACGAGCAACCTCG	1468	
Dd			
564	ThrTyxCysLeuAsnValSerLeuAlaAspThrAsnSerLeuAlaValValSerThrGln	593	
QY	ATTCTTGTTCTCTCACAGAGACCAGCTCGCTCCCTTTAAGTAGTGCAACAGTCGCCCTC---	1525	
Dd			
584	LeuIleMetPro-----ValProGlyIleLeuLeuLeuThrGlyGlnGluAlaGlyLeuGly	601	
QY	-----ATCTCCGTCGCTGCTGGCCATATTTGTCACTGTGATC---TCCTCTCTTG	1573	
Dd	602 GlnValProLeuLeuValGlyIleLeuLeuValLeuMetAlaValValLeuAlaSerLeu	621	
QY	1574 GTGTACAAAAAA	1585	
Dd	622 IleTyrArgArg	625	

## RESULT 5

A:Accession: A49179  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 13-Dec-1993 #sequence\_revision 18-Nov-1994 #text\_change 08-Nov-1996  
 C:Accession: A49179; I45861  
 R:Kim, R.Y.; Wistow, G.J.  
 Exp. Eye Res. 55, 657-662, 1992  
 A:Title: The cdna RPE1 and monoclonal antibody HMB-50 define gene products preferential  
 A:Reference number: A49179; PMID:1478275  
 A:Accession: A49179  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-491 <XIM>  
 A:Experimental source: retinal pigment epithelium  
 A:Note: sequence extracted from NCBI backbone (NCBIN:122438, NCBIPI:122439)  
 C:Genetics:  
 A:Gene: RPE1

Alignment Scores:	2.38e-29	Length:	491
Pred. No.:	Score:	Matches:	124
	436.50	Conservative:	63
Percent Similarity:	40.21%	Mismatches:	123
Best Local Similarity:	26.67%	Indels:	155
Query Match:	9.00%	Gaps:	14
DB:	2		

US-10-039-272-1 (1-2661) x A49179 (1-491)

QY		603	CAGTATTTCACGAAATTTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACACGCCAATGTG	6662
			:          :    :    :    :	
Dd		1	GlnTyrTrpGlnValLeuGlyGlyProValSerGlyLeuSerIleGlyThrAspLysAla	20
			:          :    :    :    :	
QY		663	ACACTTGGCGCTCAACTCATUGGAGTGAAGTGTCTACAGAAGACATGGA---CGGCGCATAT	719
			:          :    :    :    :	
Dd		21	MetLeuGlyThrTyrAsnMetGluValThrValTyrHisArgArgGlySerGlnSerTyr	40
			:          :    :    :    :	
QY		720	GTTCCCATCGCACAAAGTCAAAGATGTGTACGTGGTAACAGATCAGATTCCCTGTGTTTGTG	779
			:          :    :    :    :	
Dd		41	ValProLeuAlaHisSerSerSerAlaPheThrIleThrAspGlnValProPheSerVal	60
			:          :    :    :    :	
QY		780	ACTATGTTCCAG-----AAGAACGATCGAAATTCATCCGACCAAAACC	821
			:    :    :    :    :    :    :	
Dd		61	SerValSerGlnLeuGlnAlaLeuAspGlyArgAsnLysArg--PheLeuArgLysGln-	79
			:          :    :    :    :	
QY		822	TTCGCAAGATCTCCCCATTATGTTTGATTCCTCGATTCANGATCCTAGCCACTTCCTCA	881
			:          :    :    :    :	
Dd		80	-----ProLeuThrPheAlaLeuGlnLeuHisAspProSerGlyTyrLeuA	95
			:          :    :    :    :	
QY		882	ATTATTCTACCAANTAACTACAAGTGGAGCTTCGGGGATAATACTGCCTGTTGTTTCCA	941
			:          :    :    :    :	
Dd		95	laGLValaAspLysSerTyrThrTrpAspPheGlyAspSerThrGlyThrLeuIleSerA	115
			:          :    :    :    :	

QY	942	CCAATCATACTGTGTAATCACAGTATGTGCTCAATGGAAACCTTCAGGCTTAACTCACTG	1001
Db	115	rgAlaLeuThrValThrHisThrIyrLeuGluSerGlyProValThrAlaGlnValValL	135
QY	1002	TCGAAAGCTGCACACCA-----GGACCTTGTCCGCCACCCGCCACCAACACCA	1049
Db	135	euglnAlaLalLeProLeuThrSerCysGlySerProValProGlyThrThrAspA	155
QY	1050	GA-----	1051
Db	155	rgHisValThrThrAlaGluAlaProGlyThrThrAlaGlyGlnValProThrThrGluV	175
QY	1051	-----	1051
Db	175	alMetGlyThrThrProGlyGlnValProThrAlaGluAlaProGlyThrThrValGlyT	195
QY	1051	-----	1051
Db	195	rpValProThrThrGluAspValGlyThrThrProGluGlnValAlaThrSerLysValL	215
QY	1051	-----	1051
Db	215	euSerThrThrProValGluMetProThrAlaLysAlaThrGlyArgThrProGluValS	235
QY	1052	-----CCTTCAAAACCCACCCCTTCTTAGGA-----	1078
Db	235	erThrThrGluProSerGlyThrThrValThrGlnGlyThrThrProGluLeuValGluT	255
QY	1079	-----CCTGCTGGTGACAAC-----C	1094
Db	255	hrThrAlaGlyGluValSerThrProGluProAlaGlySerAsnThrSerSerPheMetP	275
QY	1095	CCCTGGAG-----CTGAGTAGAGATTCCTGATGAA-----	1123
Db	275	roThrGluGlyThrAlaGlySerLeuSerProLeuProAspThrAlaThrLeuValL	295
QY	1124	-----AATGCCAGATTAAACAGATATGCCCACTTTCAGGCCA	1160
Db	295	euGluLysArgGlnAlaProLeuAspCysValLeuTyrArgTyrGlySerPheSerLeuT	315
QY	1161	CCATCAAAATTGTAGAGGGAATCTTAGAGTTAAACATCATCCAGATGACAGAGCTCCTGA	1220
Db	315	hrLeuAspIleValSer-----lleGluSerAlaGluIleLeuG	328
QY	1221	TGCCGTGCCATGGCCTGAAAGCTCCCTAATAGACTTTGTGTGACCTGCCAAGGAGCA	1280
Db	328	lnAlaValSerSerSerGluGlyAspAlaPheGluLeuThrValSerCysGlnGlyGlyL	348
QY	1281	TTCCACGGAGGTCTGTACCATCATTTCTGACCCCACTCCGAGATCACCCAGAACACAG	1340
Db	348	euProLysGluAlaCysMetAspIleSerSerProGlyCysGlnLeuProAlaGlnArgL	368
QY	1341	TCCTGACCCCTGTGGATGTGGATGAGATGTGTCTGTGACTGTGACAGCAACCTTC---A	1397
Db	368	euCysGlnProValProProSerProAlaCysGlnLeuValLeuHisGlnValLeuLysG	398
QY	1398	ATGGGTCTGGACGTACTGTGTGAACCTCACTACCCCTGGGGGATGACACAGCCTGGCTCTCA	1457
Db	388	lyGlySerGlyThrTyrCysLeuAsnValSerLeuAlaAspAlaAsnSerLeuAlaMetV	408
QY	1458	CGACGACCCCTGATTTCTGTTCTCGACAGACCCAGCCTCGCCTTTAAGGATGGCAACA	1517
Db	408	alSerThrGlnLeuValMetProGlyGlnGlu-----AlaGlyLeuArgGlnAlaPro-	425
QY	1518	GTGCCCTGATCTCGCTTGGCTGTGTGGCCATATTGTGACT---GTGATCTCCCTCTTGG	1574
Db	426	-----LeuPheValGlyIleLeuLeuValLeuThrAlaLeuLeuLeuAlaSerLeuL	443
QY	1575	TGTACAAAAA	1585
Db	443	leTyrArgArg	446

T07907  
hydroxyproline-rich glycoprotein GAS28 precursor - Chlamydomonas reinhardtii  
C:Species: Chlamydomonas reinhardtii  
C:Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 08-Oct-1999  
C:Accession: T07907  
R:Rodriguez-Martinez, H.; Haring, M.A.; Von Gromof, E.; Beck, C.F.  
submitted to the EMBL Data Library, July 1997  
A:Reference number: Z16207  
A:Accession: T07907  
A:Status: preliminary; translated from GB/EMBL/DDbJ  
A:Molecule type: mRNA  
A:Residues: 1-446 <RD>  
A:Cross-references: EMBL:AF015883; NID:G2384727; PIDN:AA69862.1; PID:G2384728  
A:Experimental source: gametes  
C:Genetics:  
F:1-29/Domain: signal sequence #status predicted <SIG>  
F:30-446/Product: hydroxyproline-rich glycoprotein GAS28 #status predicted <MAT>

Alignment Scores:			
Pred. No.:	0.00342	Length:	446
Score:	135.00	Matches:	85
Percent Similarity:	28.13%	Conservative:	25
Best Local Similarity:	21.74%	Mismatches:	123
Query Match:	2.78%	Indels:	158
DB:	2	Gaps:	13

US-10-039-272-1 (1-2661) x T07907 (1-446)

QY	429	GCCTGGTTTCTGCTCATCCATATGTTTACAAC	TGTCAGACAGATGTCAGACGACAGTGAC	488
Db	25	AlaProAlaSerAlaAspSerGluTyrPheGlyPhe	ArgGlyGlyGlyArgValLeuHi	44
QY	489	GGGGAAATGGCACCGCCGCAAGCCATCATAC	AGCTTCCCTGTAGTGGAAACCTTTTCCT	548
Db	44	SGlnLysLysAlaLysAlaLysAlaAlaProThr	ProSer	57
QY	549	CACCACCCGGATGGAGAGATGGAAATTCAT	PCTACGTCTTCCACACACTTGGTCAGTAT	608
Db	58	-----	ProAlaAlaThrSer-ProAlaProLysLeuValP	69
QY	609	TTCCGAAATGGGACGATGTTTCAGTGAGAGT	TCTGTGAACACAGC-----	655
Db	69	heProProLeuCysAsn-----	CysLysArgAsnProArgLysSerP	83
QY	656	-----CAATGTGACACTTGGGCTCAACTCA	TGAGAGTGACTGTCTCAGAGA	704
Db	83	roPheArgMetSerLeuAspAspAlaAlaSer	AlaGlyGlyValTyrCysPheThrIleG	103
QY	705	CATGACGGGCATATGTTCCCATCGCACAG	TGGAAGATGTACGTGGTAAACACGAC	764
Db	103	lnAsnValGlyGlyCysAsp-----	ProThrGlnLysCysCysAspGlyThrGlnGly	121
QY	765	ATTCTGTGTTTGTGACTATGTTCAGAGA	ACGATCGAAATTCATCGACGAAACCTTC	824
Db	121	alSer-----	-----	122
QY	825	COAAAGATCTCCCAATTATGTTGATGCTCT	CAATTCATGATCTTCAGCACCTTCCTCAATT	884
Db	123	-----LysValGlnLeuAspValValAlaSer	CysLysAspSerValArgAsnV	139
QY	885	ATTCTACCAATTAACATCAAGTGGACCTTC	GGGGATATACTGGCTGTGTTGTT-----T	938
Db	139	alThrValAspGlyLysLysTrpSerTyrGlu	PheAsnThrAlaLeuSerValIleArgM	159
QY	939	CCACCAATCATACTGTGAATCACAGTATG	CTCAATGGAAACCTTCAGCCTTAACCTCA	998
Db	159	etThrAsnLeuGlyLysThrAlaSerThrAla	AlaGlyThrGluValCysLeuAlaLeuA	179
QY	999	CTGTGAAGCTGCAGCACCA	-----	1021
Db	179	laProLysSerGlnCysProGlyLeuAlaGln	LeuCysSerAlaGlyAlaGlyLeuCysL	199

QY	177	GAGCACAAATCAATTAAAAT-----GGCTGG	200
Db	204	GlupheAsnTyrLeuAenThrAlaArgThrLeuGluLeuTyrGlyValGluPheHisTyr	223
QY	201	TCTTCTCATGAATAACTCGAATGAAAACTCTACCCAGTCGTGGAAAGCGGGAGACATG	260
Db	224	AlaArgaspGlnSerAen---AsnGluIleMetileGlyValMetSerGlyGlyIleLeu	242
QY	261	AGGTGGAAAAACTCCTGGAAGGAGGCCGTGTGCAGCGGTCTCTGACCAGTGACTCACCA	320
Db	243	IleTyrLysAsn-----ArgValargMetAsnThrPhePro	254
QY	321	GCCCTCTGTGGGCTCAAATATAACAATTTCCGGTGAACCTGATATTCCCTAGATGCCAAAAG	380
Db	255	TrpLeuLyslleValLysIleSerPheLysCysLysGlnPhePheIleGlnLeuArgLys	274
QY	381	GAAGATGCCAATGGC-----AACATAGTCTATGAGAAGAAGCTGC	419
Db	275	GluleuHisGluSerArgLuuThrLeuLeuGlyPheAsnMetValAsnTyrArgAlaCys	294
QY	420	AGAAATGAGGCTGTTTTATCTGCTGATCCATATGTTTACAACCTGGACAGATGCTCAGAG	479
Db	295	LysAsn-----LeuTrpLysAlaCysValGlu	303
QY	480	GACAGTGACGGGGAAATGGCACCGGCCAAAGCCATCAATACGTCCTTCCTGTAGTGGAAA	539
Db	304	-----HisHisThrPhePheArgLeuAspArg	312
QY	540	CCTTTTCTCACCACCCGGATGGAGAGATGGAAATTCATCTACGCTCTCCACACACTT	599
Db	313	ProLeuProGlnLys-----AsnPhePheAlaHisTyrPheThrLeu	327
QY	600	GTCAGTATTTCCAGAMAATTCGGCAGATGTTCACTG-----	635
Db	328	GlySerLysPheArgTyrCysGlyArgThrGluValGlnSerValGlnTyrGlyLysGlu	347
QY	636	-----AGAGTTTCTGTGAACACAGCCAATGTGACACTTGGCGCCTCAACTC	680
Db	348	LysAlaAsnLysAspArgValPheAlaArgSerProSerLysProLeuAlaArgLysLeu	367
QY	681	ATGGAAAGTGACTGTCTACAGAAGACATGCAGCGGCATATGTTCCCATCGCAAGATGAAA	740
Db	368	MetAspTrpGluValValSerArgAsn-----SerIleSer	379
QY	741	GATGTGTACGTGTAAACAGATCATCATCTCTGTGTTGTGACTATGTTCCAGAAGAACGAT	800
Db	380	AspAspArgLeuGluThrGlnSerLeuProSerArgSerProGlyThrProAsnHis	399
QY	801	CGAAATTCATCGA-----CGAAACCTTCCAAAGATCTCCCCATTATG	844
Db	400	ArgAsnSerThrPheThrGlnGluGlyThrArgLeuArgProSerSerValGlyHisLeu	419
QY	845	TTTGTATGTCCTGATTTCAT--GATCCTAGCCAC--TTCCCTCAATTATTCTACCATTAAC	898
Db	420	ValAspHisMetValHisThrSerProSerGluValPheValAsnGlnArg-----	436
QY	899	TACAAAGTGGAGCTTCGGGGATAAATAGTCGCCTGTTGTTTCCOCAAATCACTGTGTAAT	958
Db	437	-----SerProSerSerThrGlnAla	443
QY	959	CACACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCACACCA	1018
Db	444	AnSerIleValLeuGluSerSerProSer-----GlnGluThrPro	457
QY	1019	GGACCTTGTCCGCCACCGCCACACCCACCA-----	1051
Db	458	GlyAspGlyLysProProhalaLeuProProLysGlnSerLysLysAsnSerTrpAsnGln	477
QY	1051	-----	1051
Db	478	IleHisTyrSerHisSerGlnAspLeuGluSerHisIleAsnGluThrPheAspile	497

Qy	1052	-----CCTCAAAACCACCCCTCTCTTTAGGACCTGCTGGTGACAAACCCCTGGAG	110
Db	498	ProSerProGluLysProThrProAsnGlyGlyLeProHisAspAsnLeuValLeu	517
Qy	1103	CTGAGTAGGATTCTCATGAAAACTGCCAGATTAAACAGATATGCC	1147
Db	518	IleArgMetLysProAspGluAsn-----GlyArgPheGly	529
RESULT 8			
T18518			
apolipoprotein(a) - western European hedgehog (fragment)			
C:Species: Erinaceus europaeus (western European hedgehog)			
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 07-Dec-1999			
C:Accession: T18518			
R:Lawn, R.M.; Boonmark, N.W.; Schwartz, K.; Lindahl, G.E.; Wade, D.P.; Byrne, C.			
J. Biol. Chem. 270, 24004-24009, 1995			
A:Title: The recurring evolution of Lp(a): Insights from cloning of hedgehog ap			
A:Reference number: 146259; MUID:96025778; PMID:7592597			
A:Accession: T18518			
A>Status: preliminary; translated from GB/EMBL/DBD			
A:Molecule type: mRNA			
A:Residues: 1-2869 <LAW>			
A:Cross-references: EMBL:U33170; NID:g1046358; PID:g1046359; PIDN:AAC48522.1			
A:Experimental source: liver			
C:Comment: The lipoprotein Lp(a), a major inherited risk factor for atheroscler			
ent apolipoprotein(a).			
Alignment Scores:			
Pred. No.: 0.162 Length: 2869			
Score: 118.00 Matches: 141			
Percent Similarity: 29.39% Conservative: 70			
Best Local Similarity: 19.64% Mismatches: 248			
Query Match: 2.43% Indels: 259			
DB: 2 Gaps: 31			
US-10-039-272-1 (1-2661) x T18518 (1-2869)			
Qy	6	CGAGGGCCGACGAGAAATGAAGTTAACCTTGGTGCCTCGTCGTCGAGAATTCAGCATGGAA	65
Db	410	ArgAsnProAspGlyGluLeuAlaProThrCys-----TyrThr-As	424
Qy	66	TGTCTCTACTATTCTCTGGGATTTCTCTCTGGCTGCAAGATTGCCAC---	122
Db	424	n---SerAlaValArgTrpGluTyrCysSer---IleProAlaCysGluSerLeuSerPr	442
Qy	123	GCCAAACGATTTCATGATGCTCT---GGGCAATGAAGACCTCTCTGCTTACATGAGGAG	179
Db	442	oProThrGluProMetVal-ValProGlyGln-----CysLeuGluGlyThr	458
Qy	180	CACAAATCAATTAATATGGC-----TGG	200
Db	458	lyGluAsnTyrArg-GlySerValAlaValThrValSerGlyHisThrCysGlnArgTrp	477
Qy	201	-----TCTTCTGATGAAAATGACTGGAATGAAAACCTCTACCCAGTGTGGAAGCGG	251
Db	478	ArgGluGlnSerProHisSerHisSerTrpThrProGluAsnTyrProThrLysAsnLeu	497
Qy	252	GGAGCATAGGTGGAAAACTCTCGGAGGAGGCCCTGTGCGAG-----	296
Db	498	AspGlyAsnTyrCysArgAsnThr-----GlyGlyGluValAlaProTrpCysTyrThr	515
Qy	297	-----GGGTCCTCGACGACGACTCACACGCCCTC	326
Db	516	ThrAsnSerAlaValArgTrpGluTyrCysSerIleProAlaCysGluSerProThr---	534
Qy	327	GTGGGCTCAATATAACATTTCGGGTGAACCTGATATTCCTAGATGCCAAAGAAAGAT	386
Db	535	-----ProProThrGluGlnGlnVal	541
Qy	387	GCCATGGCAACATAGTC-----TATGAGAGAACTGCAGAAATGAGCTGTTTATCT	440
Db	542	ValProGlyGlnCysLeuGluGlyThrGlyGluAsnTyrArgGlySerValAlaValThr	561

QY	441	GCTGATCCATATGTT---TACAACTGGACAGCATGGTTCAGAGCAGCATGACGGGAAAAT	497
Db	562	AlaSerGlyHisThrCysGlnGlnTrpArgAlaGlnSerProHisSer-----HisSer	579
QY	498	GGCACGGGCCAAAGCCATCAATACGTCCTCCCTGATGGAAAACCTTTTCCTCACACACCC	557
Db	580	ArgThrProGluAsnTyPro-ThrLysAsnLeuValGlyAsn---TyrCysArgAsnPr	598
QY	558	GGATGGAGAAGATGGAATTTCACTACGCTTTCCACACACTTGGTCAGTATTTCAGAAA	617
Db	598	oAspGlyGluIleAlaProTrpCysTyThrThrAsnSerAlaVal-----Ar	614
QY	618	TTGGACGATGTTAGTGAGAGTTTCTGTGAACACAGACCAATGTGACATTGGGCTCAA	677
Db	614	gTrpGlu-TyrCysSerIleProAlaCysGluSerSerProProThrGluProMetV	634
QY	678	CTCATGGAAGTGACTGTCTACAGAAGACA-----TGGACGGGCATATGTTCCCATC	728
Db	634	alValProGlyGlnCysLeuGluGlyThrGlyGluAsnTyArgGlySerValAlaValT	654
QY	729	GCACAAGTGAAGA---TGCTGACTGTGTAACAGATCAGATTCCTGTGTTGTGACATATG	785
Db	654	hrValSerGluHisThrCysGlnArgTrpArgGluGlnSerProHisSerHisSerTrpT	674
QY	786	TTCCAGAAGACGATCGAAATTCATCCGACGAAACCTTCCCAAAGATCTCCCCATTATGT	845
Db	674	hrProGlu-----	676
QY	846	TTGATGTCCTGATTCATGATCCTAGCCATCTCTCAATTATTCACCATTAACCTACAGT	905
Db	677	-----AsnTyProThrLysAsnLeuAspG	685
QY	906	GGAGCTTCGGGGATAATACTAGCCCTGTTGTT-----TCCACCAATCATA	950
Db	685	lyAsnTyrcysArgAsnThrGlyGlyGluValAlaProTrpCysTyThrThrAsnSerA	705
QY	951	CTGTGATTCACAGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTG	1010
Db	705	laVala-gTrpGluTy-----CysS	712
QY	1011	CAGCACGAGGACCTTGTCGCGCACCGGCACCAACCACCCAGACCTTCAAACCCACCCCTT	1070
Db	712	erIleProAlaCysGluSerProThrProThrGluGlnGlnValValProGlyGlnC	732
QY	1071	CTTTAGGACCTGCTGGTGACAAAC-----	1093
Db	732	ysLeuGluGlyThrGlyGluAsnTyArgGlySerValAlaValThrAlaSerGlyHisT	752
QY	1094	-----CCCTGGAGCTGAGTAGGATTCCTGATGATAAACT	1127
Db	752	hrCysGlnGlnTrpArgAlaGlnSerProHisSerHisSerArgThrProGluAsnTyPr	772
QY	1128	GCCAGATTACAGATATGGCCACTTTCAAGCCACCATCACAACTGTAGAGGGAACTCTTAG	1187
Db	772	roThrLysAsnLeuValGlyAsnTyrcysArgAsn-----	783
QY	1188	AGGTTAATCATCATCCAGATCACACAGCTGCTGATCGCGGTGCCATGGCTGAAGCTCCC	1247
Db	784	-----ProAspGlyGluI	788
QY	1248	TAATAGACTTTGTGCTGACTGTGCCAAGGAGCATTCACCGAGGCTCTGTACCATCATTTT	1307
Db	788	leAlaProTrpCysTyThrThrAsnSerAlaValArgTrpGluTyrcysSerIle-----	806
QY	1308	CTGACCCCACTCGGAGATCACCAGAACACAGTCTGCACCCCTGTGGATGTGGATGAGA	1367
Db	807	-----ProAlaCysGlu---SerLeuSerProThrGluProMetValValProGlyG	824
QY	1368	TGTGTCGTGCTGACTGTGAGACGAACCTTCATGGTGCT-----	1405
Db	824	lnCysLeuGluGlyThrGlyGluAsnTyArgGlySerValAlaValThrValSerGlyH	844
QY	1405	-----	1405

Db 844 isThrCysGlnArgTrpArgGluGlnSerProHisSerHisSerTrpThrProGluAsnT 864  
 QY 1406 -----GGGACGTACTGTGTGAACCTCACCTCGGGGAT----- 1438  
 Db 864 yProThrLysAsnLeuAspGlyAsnTrpCysArgAsn---ThrGlyGlyGluValalAp 883  
 QY 1439 -----GACACAAGCCTGCTCTCACGACACCCCTGATTTCTTCTCTCTGAC--A 1484  
 Db 883 roTrpCysTyThrThrAsnSerAlaValArgTrpGluTrpCysSerIleProAlaCysG 903  
 QY 1485 GAGACCCAGCCTCGCCCTTAAAGTATGCAACACAGTGCCTGATCTCCGTTG----- 1535  
 Db 903 luSerProThrProThrGluGln-GlnValValProAspGlnCysLeuGluGlyThr 922  
 QY 1536 -----GCTGCTTGGCCATATTGT----- 1554  
 Db 923 GlyGluAsnTrpArgGlySerValAlaValThrAlaSer-GlyHisThrCysGlnGlnTr 942  
 QY 1554 ----- 1554  
 Db 942 pArgAlaGlnSerProHisSerHisSerArgThrProGluAsnTrpProThrLysAsnLe 962  
 QY 1555 -----CACGTGATCTCCCTCTTGGTGACAAAAACACAGGATACACCCAA 1604  
 Db 962 uValGlyAsnTrpCysArgAsnProAspGlyGluIleAlaProTrpCysTrpThrAs 982  
 QY 1605 TAGAAAATAGTCTCTGGGAATGTGTGACAGCAAGGCGCTGAGTGTCTTCTCAACCGTG 1664  
 Db 982 nSerAlaValArgTrpGluTrpCysSerIleProAlaCysGluSerLeuSerProThr 1002  
 QY 1665 CAAAGCGGTGTTCTCCCGGAAACACGAAAGATGCC 1704  
 Db 1002 rGluProMetValProGlyGlnCysLeuGluGlyThr 1015

RESULT 9  
 S21766  
 dihydroliipoamide S-acetyltransferase (EC 2.3.1.12) - rat (fragment)  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 22-Nov-1993 #sequence\_revision 01-Feb-1999 #ext\_change 05-May-2000  
 J:Accession: S21766  
 R:Matuda, S.; Nakano, K.; Ohta, S.; Shimura, M.; Yamanaka, T.; Nakagawa, S.; T  
 Biochim. Biophys. Acta 1131, 114-118, 1992  
 A:Title: Molecular cloning of dihydroliipoamide acetyltransferase of the rat pyru  
 S:  
 A:Reference number: S21766; MUID:92256482; PMID:1581353  
 A:Accession: S21766  
 A:Molecule type: mRNA  
 A:Residues: 1-555 <MAT>  
 A:Experimental source: heart  
 C:Genetics:  
 A:Genome: nuclear  
 C:Complex: component E2 of pyruvate dehydrogenase complex  
 C:Function:  
 A:Description: catalyzes conversion of acetyl-CoA and dihydroliipoamide to S-acet  
 A:Pathway: pyruvate metabolism  
 C:Superfamily: dihydroliipoamide acetyltransferase; lipoyl/biotin-binding homolog  
 C:Keywords: acetyl-CoA; acyltransferase; cardiac muscle; coenzyme A; heart; lipo  
 F:7-81/Domain: lipoyl/biotin-binding homology <LPB1>  
 F:133-207/Domain: lipoyl/biotin-binding homology <LPB1>  
 F:254-307/Domain: component E3 binding #status predicted <EPB2>  
 F:46,172/Binding site: liipoamide (Lys) (covalent) #status predicted <EB3>  
 F:528,532/Active site: His, Asp #status predicted

Alignment Scores:  
 Pred. No.: 0.146 Length: 555  
 Score: 116.50 Matches: 78  
 Percent Similarity: 34.90% Conserves: 41  
 Best Local Similarity: 22.87% Mismatches: 129  
 Query Match: 2.40% Indels: 93  
 DB: 2 Gaps: 15

US-10-039-272-1 (1-2661) x S21766 (1-555)





Db	1445	-----PheAsp-----	SerHisAlaGluCysIleAlaT	1454
QY	876	TCCTCAATTATTCTACCAATT		
Db	1454	YrLeuSerThrCysThrValAlaArgThrGlyCysGlnAlaArgAlaThrCysAlaS		
QY	897	ACTACAG		
Db	1474	erTyrLysSerSerGlnGlnCysLysPheAsnSerThrGlyGlyLysCysPheTTPAsnP		
QY	912	TCGGGAT		
Db	1494	roAsnAsnLysThrCysValAspLeuAsnCysGlyAsnIleGluAlaThrThrThrTyra		
QY	954	TGAATCACACGATGTGCTCAATGGAACTTCAGCCTTAACCTCACTGTGAAA		
Db	1514	spThrHisAlaGluCysValAlaValAspThrGlnLeuLeuCysThrValArgAlaThrA		
QY	1007	-----GCTGCACACACGAGCACTTGTCCGCCACCGCCACCCACCCAGACCTTCAAAAC		
Db	1534	snGlyAlaAlaValProGly		
QY	1062	CCACCCCTCTTTAGACACCTGCTGGTGACACCCCTGGAGCTGAGTAGGATTCCTGATG		
Db	1541	--CysMetAlaArgGlyAlaCysSerSerTyrSerIleGlu		
QY	1122	AAAACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATCACTAATTGTAGAGGAA		
Db	1554	spGlnCysLysThrAsnProSerGly		
QY	1182	TCTTAGAGTTAAACATCATCCAGATGACAGAGCTCCTGATGCCGGTGCATGGCCTGAAA		
Db	1564	alCysValTrpAsn		
QY	1242	GCTCCCTAATAGACTTTGCTGTACC		
Db	1580	erCysThrThrAlaProThrAlaThrHisAlaAspCysAspSerTyrPheSerT		
QY	1281	TTCCACCGAGGTCTGTACCATCATTTCTGACCCACC		
Db	1600	hrAlaThrIleLysCysThrValAlaAlaThrProAspThrAsnGlyGlyAlaAlaValL		
QY	1319	-----TGCAGATACCCAGAAACACAGCTCAGCCCTGTGGATGTGATGAGATGT		
Db	1620	euGlyGlyCysGln		
QY	1371	GTCTGCTGACTGTGAGAGCAACCTCAATGGGTCTGGAGCGTACTGT		
Db	1638	ys		
QY	1428	CCCTGGGGGATGACACAGAGCTGCTCTCAGACACCCCTGATTTCTGTCCTGACAGAG		
Db	1652	hrGlnCysAlaAspLysSerCysAlaThrAlaProAlaThrThrAspTyrAspAspAsnA		
QY	1488	AC		
Db	1672	sp		

RESULT 11  
T02831  
AAA protein L4171.3 [imported] - Leishmania major (strain Friedlin)  
C:Species: Leishmania major  
C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 19-May-2000  
C:Accession: C81460; T02831  
R:Myler, P. J.; Audleman, L.; deVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magnus, C.;  
Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999  
A:Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c  
A:Reference number: A81455; MUID:99178987; PMID:10077609  
A:Accession: C81460  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1541 <PYL>  
A:Cross-references: GB:AE001274; NID:g3264850; PIN:AAC24654.1; PID:g2266909; GSPDB:GNOC

```

A;Experimental source: strain MHOM/IL/81/Friedlin
C;Genetics:
A;Gene: L4171.3
A;Map position: 1

Alignment Scores:
Pred. No.      0.377      Length:      1541
Score:         113.00     Matches:     46
Percent Similarity: 37.58%    Conservative: 16
Best Local Similarity: 27.88%  Mismatches:  60
Query Match:    2.33%     Indels:     43
DB:             2        Gaps:       6

```

US-10-039-272-1 (1-2661) x T02831 (1-1541)

QY	1022	CTTGTCCGCCACCGCCACCCACCCAGACCTTCAAAACCCACCCCTTCTTTAGGACCT	1081
Db	156	ProProProProProProProProProProProProMetMetMetProProLeuAsnGly-Me	175
QY	1082	GCTGGTGACAACCCCTGGAGCTGAGTAGGATTCCTGATGAAAACCTGCCAGATTAAACAGA	1141
Db	175	tilePhe-GlnProProGlyThr-----	182
QY	1142	TATGGCCATTTCAAGCCACCATCAAAATGTAGAGGAATCTTAGAGGTTAAACATCATC	1201
Db	183	-----ProGlyMetProHisHisHis-----	192
QY	1202	CAGATGACAGACGCTCGTGATGCGGGTGCCATGGCCCTGAAAGCTCCCTAATAGACTTTGTGCT	1261
Db	192	ro-----AlaGlyMetAlaThrProGln-----	204
QY	1262	GTGACCTGCTCAAGGGAGACATTCACCGGAGGCTGTGPACCATCATTTCTGACCCACCT--	1319
Db	204	erGlnHisGlnHisGlnHisGlnHisGly--MetMetProMetTrpMetProProPro	223
QY	1320	-----GGAGATCACCACGACACAGCTCTCAGCCCTGTGGATGTGATGAGATGTGT	1372
Db	223	oThrGlnThrGlnGlnHisGlnGlnGlnGlnProGlyMetMetMetProMetAsnSerMe	243
QY	1373	CTGCTGACTGTGACAGACCACTTCAATGGTCTGGAGCTACTGTGTGAACCTCACCTGTG	1432
Db	243	tMetProMetSerMetProMetSerMetProMetGly-AlaAlaSerSerProLeuPro	263
QY	1433	GGGGATGACAAAGCGTGGCTCTCACGAGCACCTCGATTCTGTCTCTGACAGACCA	1492
Db	263	roGlyGlnHisSerSerGlyProGlnGlnHisPro-----	279
QY	1493	GCCTC	1497
Db	279	lyMet	280

RESULT 12  
155976  
dihydrolipoamide S-acetyltransferase (EC 2.3.1.12), liver - rat (fragment)  
N:Alternate names: primary biliary cirrhosis autoantigen  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 05-May-2000  
C:Accession: I55976; I71932  
E:Gershwin, M.E.; Mackay, I.R.; Sturgess, A.; Coppel, R.L.  
J. Immunol. 138, 3525-3531, 1987  
A:Title: Identification and specificity of a cDNA encoding the 70 kd mitochondrial  
A:Reference number: I55976; MUID:87196380; PMID:3571977  
A:Accession: I55976  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-457 <RES>  
A:Cross-references: GB:D00092; NID:g220817; PID:AAA41813.1; PID:g220818; GB:M16  
A:Experimental source: liver  
C:Genetics:  
A:Genome: nuclear  
C:Complex: component E2 of pyruvate dehydrogenase complex  
C:Function:  
A:Description: catalyzes conversion of acetyl-CoA and dihydrolipoamide to S-acetyl-

A: Pathway: pyruvate metabolism  
C: Superfamily: dihydrolipoamide acetyltransferase, lipoyl/biotin-binding homology  
F: 48-122/Domain: lipoyl/biotin-binding homology <LPB2>  
F: 169-222/Domain: component E3 binding #status predicted <E3B>

Alignment Scores:  
Pred. No.: 0.309 Length: 457  
Score: 112.50 Matches: 75  
Percent Similarity: 34.97% Conservative: 39  
Best Local Similarity: 23.01% Mismatches: 121  
Query Match: 2.32% Indels: 91  
DB: 14 Gaps: 14

US-10-039-272-1 (1-2661) x I55976 (1-457)

QY 957 ATCACAGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTGTCAAGCTGCAGCAC 1016  
DB 128 ILAALAPheAlaAspTyrArgProThrGluValThrSerLeu-LysProGlnAlap 147  
QY 1017 CAGGACCTTGTGCGCCACCG-----CCAGCACCCAGACCTTCAAAACCCA 1064  
DB 147 roProValProProValAlaAlaValProProlleProGlnProLeuAlaPro 167  
QY 1065 CCCCTCTTTAGACCTGCTGGTGAC-----AACCCCTCGAGCTGA 1106  
DB 167 hrProSerAlaAlaProAlaGlyProLysGlyArgValPheValSerProLeuAlaLysL 187  
QY 1107 GTAGGATTCCTGTAAGAACTGCCAGATTAAACAGATTAGCCACTTTCAAGCCACCATCA 1166  
DB 187 ysLeuAlaAlaGluLysGlyLeuAspLeuThr-----GlnValLysGlyT 202  
QY 1167 CAATTGTAGAGGA---ATCTTAGAGTTTAACATCATCCAGATCAGACAGCTCTGTATGC 1223  
DB 202 hrGlyProGluGlyArgIleIleLysLysAspIle-----AspSerPheValP 218  
QY 1224 CG-----GTGCCAT 1232  
DB 218 roThrLysAlaAlaProAlaAlaAlaAlaAlaProGlyProArgValAlaProT 238  
QY 1233 GGCCTTGAAGCTCCCTAATAGACTTTGTGCTGACCTGCCMAGGGAGCATCCCGAGGAGG 1292  
DB 238 hrProAlaGlyValPheIleAspIleProIle-----SerAsnI 251  
QY 1293 TCTGTACATCATTTGTGACCCCGAGATCACCAGAACACAGCTGTGAGCCCT- 1351  
DB 251 leArgArgValIleAlaGln-----ArgLeuMetGlnSerLysGlnThrIleProH 268  
QY 1352 -----GTGGATGTGATGAGATGTGCTGCTGACTGTGAGACGAACTTCA 1397  
DB 268 isTyrTyrLeuSerValAspValAsnMetGlyGluValLeuLeuValArgLysGluLeuA 288  
QY 1398 AT-----GGGTCTGGGAGCTACTGTGTGAACCTCACCTCGGGGATGACACAA 1445  
DB 288 snLysMetLeuGluGlyLysGlyLysIleSerValAsnAspPheIleIleLysAlaSerA 308  
QY 1446 GCCTGGCTCTCACGAGCACCCTGATTTCTGCTCAGACAGACCCAGCC-----T 1496  
DB 308 laLeuAla-----CysLeuLysValProGluAlaAsnSerSerTrpMetAspT 324  
QY 1497 CGCCTTTAAGGATGCCAACAGTGCCTGATCTCGCTT----- 1534  
DB 324 hrValIleArgGlnAsnHisValValAspValSerValAlaValSerThrProAlaGlyL 344  
QY 1535 -----GGCTGTGTGGCCATATTGTCACCTG 1559  
DB 344 euIleThrProIleValPheAsnAlaHisIleLysGlyLeuGluThrIleAlaSerAspV 364  
QY 1560 TGATCTCCCTCTGTGTGTAACAAACACAGGAATACACCCCAATAGAAATAGTCCCTG 1619  
DB 364 alValSerLeuAlaSerLysAlaArgGluGlyLysLeuGlnProHisGluPheGlnGlyG 384  
QY 1620 GGAATGTGTGTCAGAAC-----AAAGCCCTGAGTGTCTTTCTCA 1658

DB 384 lYThrPheThrIleSerAsnLeuGlyMetPheGlyIleLysAsnPheSerAlaIleIleA 404  
QY 1659 ACCGTGCGAAAGCCGTGTTCTTCCCGGAAACACAGAAAGGATCGCTTACTCAAAACCC 1718  
DB 404 snProProGlnAlaCysIleLeuAlaIleGlyAlaSerGluAspLysLeuIleProAlaA 424  
QY 1719 AAGAATTTAAAGGA 1732  
DB 424 spAsnGluLysGly 428  
RESULT 13  
T15840  
hypothetical protein C54G7.3 - Caenorhabditis elegans  
C: Species: Caenorhabditis elegans  
C: Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C: Accession: T15840  
R: Du, Z.  
submitted to the EMBL Data Library, November 1995  
A: Description: The sequence of C. elegans cosmid C54G7.  
A: Reference number: Z18416  
A: Accession: T15840  
A: Status: preliminary; translated from GB/EMBL/DBJ  
A: Molecule type: DNA  
A: Residues: 1-2946 <DUZ>  
A: Cross-references: EMBL:U40410; NID:g1065453; PID:g1065455; PIDN:AAA81392.1; CESP:C54G  
46/1; 2561/2; 2603/2; 2626/2; 2665/2; 2716/3; 2804/3; 2884/3  
A: Gene: CESP:C54G7.3  
A: Introns: 16/1; 53/2; 92/1; 160/3; 295/1; 346/1; 392/1; 440/1; 475/3; 579/1; 615/3; 1615/3; 16  
46/1; 2561/2; 2603/2; 2626/2; 2665/2; 2716/3; 2804/3; 2884/3  
Alignment Scores:  
Pred. No.: 0.597 Length: 2946  
Score: 111.50 Matches: 126  
Percent Similarity: 30.64% Conservative: 60  
Best Local Similarity: 20.76% Mismatches: 201  
Query Match: 2.30% Indels: 220  
DB: 2 Gaps: 30  
US-10-039-272-1 (1-2661) x T15840 (1-2946)

QY 472 GGTGAGGAGCAGTGCAGGGGAAATGCGACCGGCCAAAGCCATC-----ATA 519  
DB 1101 GlyPheArgAlaThrSerGlyIleCysGluProAlaIleAlaValGlyGluProCysVal 1120  
QY 520 ACGTCTTCC-----CTGATGGGAAACCTTTTCTCCTCACCACC 555  
DB 1121 ThrSerAsnGlnCysPheAspGluSerGluCysValPheGlyIleCysThrCysThrGly 1140  
QY 556 CCGGATGGGAGAGAT-----GGAAATTTTCATCTACGTCTTCCACAC 597  
DB 1141 ProAsnCysLysAspThrLysMetAlaHisProGlyGluAspCysThrSerLeuLysThr 1160  
QY 598 TTGTGTCAGTATTTCCAGAAATGGCCAGCATGTTGAGTGCAGAGTTTCTGTGAACACAGCCA 657  
DB 1161 ValCysSer-TyrAsnSerTyrCysSerLysMetSer---SerValCysGlu----- 1176  
QY 658 ATGTGACACTTTGGGCTCAACTCATGGAAGTCACTGTCTCAGACAGACATGACGCGGAT 717  
DB 1177 -CysProSerGlyMetAlaThrLysGlyThrLysCys---GluAsnThrPheGluSerI 1195  
QY 718 A-----TGTTCCCATCCGACAAAGTGAAGATGTGATGCTGTAACATCAGATTC 768  
DB 1195 eGlyLysAspCysVal-----ThrSerArgAsnCysGlnLysSerSerTyrCysAspAs 1213  
QY 769 C-----TGCTTTCTGAC-----TATGTTCC 789  
DB 1213 nGlyTyrCysValCysLysAsnGlyHisLysIleGlyGluAsnMetCysPheAsnSerPr 1233  
QY 790 AGAAGAAGCATCGAAATTC-----ATCCGACGAAACCTTCCCAAGATCTCCC 837  
DB 1233 oSerGluTyrLysSerPheSerIleLeuProPheAspLysAsnIleGlyGlnAsnThrPr 1253

```

QY      838 C-----ATTATGTTTGATGTCCT 855
Db      1253 oLeuGlnAsnThrLeuLysAsnGluPheArgGlyLeuGlnGluLeuSerAsnAspGlyLe 1273
QY      856 GATTATGAT----- 865
Db      1273 uPheHisThrSerThrLysTrpProGluIleLeuSerPheThrMetIleProProPr 1293
QY      866 -----CCTAGCCACTTCTC----- 880
Db      1293 oAspThrAsnLeuProAsnSerAsnLeuProGlnValPheSerSerPheProIleValTy 1313
QY      881 -----AATTATCTACCAATTAACACAGNGGAGCTT 912
Db      1313 rGlyAlaLysThrValAlaGluGluAsnAsnSerThrMetCysTyrIleAlaPh 1333
QY      913 CGGGGATAACTGGCTGTTTCCACCAATCATCTGTGAATCACACGATGTGCT 972
Db      1333 e-----ProGlyGluTyrCysGlyThrGlyGlnValCysLeuGlyAsnSerValCy 1350
QY      973 CAATGGAACCTTCAGCCTTAACCTCACTGTGAAGCTGACACAGGACCTTGTCCGCC 1032
Db      1350 sGluAsnGlnPheCysArgCysLeuGlnAspValAlaAlaGluAsnGlyIleCysProPr 1370
QY      1033 ACCGCCACCAACCCAGACCTTCAAAACCCACCTTCTTAGGACCTGCTGGTGACAA 1092
Db      1370 o-----GlnValAspAsnLeuArgValLeuGlyLeuG 1381
QY      1093 CCCCCCTG-----CAGCTGAG 1107
Db      1381 nProLeuGlyLysGluPheArgPheSerGluGlyLysLysIleGluMetArgArgThrSe 1401
QY      1108 TAGGATTCCTGATGAACATGCCAGATTAAACAGATATGCCACTTTCACAGCCACCATAC 1167
Db      1401 rSerLeuProLeuGluAsnCysGlnAsnGluValCysGluAsnAsnSerThrCysG 1421
QY      1168 AATTGTAGAGGAATCTTAGAGTT----- 1192
Db      1421 nSerIleLeuGlyLeuGlyArgIleCysGlnCysValGluAsnThrValLeuTrpAsnG 1441
QY      1193 -AACTATCATCAGATGACACAGCTC-----CTGATGCCGGTGCCA----- 1231
Db      1441 yAsnCysValIleValGluAspSerTyrAspLeuThrProIleAspGlyAsnCysAsp 1461
QY      1232 -----TGGCCTGAAGCTCCCTATAGACTTTGTGTCAGCTGCCACAGGAG 1278
Db      1461 uAspSerMetCysLeuSerGlySerGluCysValAspGlyLysCysLeuCysSerAsp 1481
QY      1279 C-----ATTCCACGAGGCTGTGATCATCATCTTCTGACCCC-----ACCTGGCAGAT 1326
Db      1481 yLysArgLeuIleLeuGlyIleCysValPheIleAlaLeuProGluThrSerCysGluAs 1501
QY      1327 CACCAGAACACAGTCTGACGCCCTGTGGATGTGATGAGATGTCTGCTGACTGTGAG 1386
Db      1501 nGlyGluValCysIleAsnGlySerValCysGlyAspSerAsnCysGluCysThr--G 1520
QY      1387 ACGAACCTTCATGGTCTGGACCTACTGTGTGAACCTCACCTCCCTGGGGATGACACAG 1446
Db      1520 uAsnThrTyrAsnHisAsnGlyAsn---CysValAspIleLysLeu-----AspGluSe 1537
QY      1447 CTTGCTCTCAGCAGCACCCCTGATTCTGTTCTCTGACAGACCCAGCTCGCCTTTAAG 1506
Db      1537 rIleuLeuArgGlnGluValGlyGluAspSerAspGluLysAspSerGluSerLeu-- 1556
QY      1507 GATGCCAAACAGTGCCTGATCTCCGTTGGCTGTGCTTGGCCATATTGTCACTGTGATCTC 1566
Db      1556 ----- 1556
QY      1567 CCTCTGTGTACAAAAACACAGGAATACACCCCAATAGAAATATGCTCTGGGAATGT 1626
Db      1557 -----SerGluAsnAspAsnGluGluAsnSerArgSerLe 1569
QY      1627 GGTCAAGACCAAGCCCTGAGTGTCTTTCTCAACCGTGCMAAAGCCGTTCTTCCCGGG 1686

```

## RESULT 14

KGRUGH

histidine-rich glycoprotein precursor - human

N:Alternate names: HRG

C:Species: Homo sapiens (man)

C&gt;Date: 04-Dec-1986 #sequence\_revision 04-Dec-1986 #text\_change 16-Jun-2000

C:Accession: A01287; S29669

R:Koide, T.; Foster, D.; Yoshitake, S.; Davie, E.W.

Biochemistry 25, 2220-2225, 1986

A:Title: Amino acid sequence of human histidine-rich glycoprotein derived from the nucl.

A:Reference number: A01287; PMID:86216149; PMID:3011081

A:Accession: A01287

A:Molecule type: mRNA

A:Residues: 1-525 &lt;KOI&gt;

A:Cross-references: GB:AB005803; NID:g2280513; PIDN:BAA21613.1; PID:g2280514

R:Hennis, B.; Havelaar, A.; Klufft, C.

submitted to the EMBL Data Library, October 1991

A:Description: PCR detection of a dinucleotide repeat in the human histidine-rich glyco

A:Reference number: S29669

A:Accession: S29669

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 214-247 &lt;HEN&gt;

A:Cross-references: EMBL:Z17218; NID:g32453; PIDN:CAA78925.1; PID:g32454

C:Comment: Although its physiological function is not yet known, HRG does bind heme, dy

din, and the lysine-binding site of plasminogen. On the basis of its homology with HMW

lood coagulation cascade.

C:Comment: The amino half of this protein is homologous to the first two cystatin-like

could not have inhibitory activity.

C:Comment: In addition to having a high histidine and proline content, this protein has

e-rich' region.

C:Genetics:

A:Gene: GDB:HRG

A:Cross-references: GDB:120055; OMIM:142640

A:Map position: 3q27-3q27

C:Superfamily: histidine-rich glycoprotein; cystatin homology

C:Keywords: duplication; glycoprotein; heparin binding; tandem repeat

F:1-18/Domain: signal sequence #status predicted &lt;SIG&gt;

F:19-525/Product: histidine-rich glycoprotein #status predicted &lt;MAT&gt;

F:19-131/Domain: cystatin homology &lt;CYL&gt;

F:140-246/Domain: cystatin homology &lt;CY2&gt;

F:276-321/Region: proline-rich

F:348-437/Region: histidine-rich

F:351-497/Region: proline-rich

F:63,125,344,345/Binding site: carbohydrate (Asn) #status predicted

F:78-89,105-126,218-241/Disulfide bonds: #status predicted

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

0.431  
111.00  
29.50%  
25.50%  
2.29%  
1

Length: 525  
Matches: 51  
Conservative: 8  
Mismatch: 51  
Indels: 91  
Gaps: 13

US-10-039-272-1 (1-2661) x JQ0533 (1-1874)

Qy	797	CGATCGAAATTCATCCGACGAAACCTTCCCAAGATCTCCCAATTATGTTGATGTC---	853
Db	533	ArgMetSerPheThrValAsnLeuProLeuGlnProMetTrpLeuAlaIleGly	552
Qy	854	---CTGATTTCATGATCTAGCCACTTCTCAATTAATTTCTACCATTAACTACAAGTGG	907
Db	553	AlaserLeuValProGluLeuAlaPheLeuLeuSer---	565
Qy	908	AGCTTCGGGGATAATACTGGCGTGTGTTTTCACCAATCATACTGTGAATCACACGTAT	967
Db	566	LeuSerGlyAsp-----ValAspLeuGlnThrGlnHisAspIleTyrHisHisHis	582
Qy	968	GTGCTCAATGGAACCTTCAGCCCTTAACCTCACT---	1018
Db	583	LeuHisProGluAsnPheThrLeuSerTirThrArgThrProTyrLeuAlaLeuAlaPro	602
Qy	1019	GGACCT-----	1030
Db	603	SerProPheLeuProTyrAlaHisSerProLeuProProLeuProValAsnSerSerPro	622
Qy	1031	-----CCACGGCCACACACACCAGACCTTCAAAACCCACCCCTCTCTTAGGACCT---	1081
Db	623	LeuPheProProProProLeuProProSerGlnProProLeuSerGlnGlyProAla	642
Qy	1081	-----	1081
Db	643	ThrGlnAlaProSerAlaGlnProThrProGlyGluProLeuLeuAlaProProThrThr	662
Qy	1082	-----	1090
Db	663	GluLeuLysProGluSerSerAsnProAsnAsnProAsnProSerSerSerAlaGlySer	682
Qy	1091	AACCCCTGGAGCTGAGTAGATTCCTCATGAAACTGCCAGATTAAACAGATATGGCCAC	1150
Db	683	AsnProProProLysSerSerSerAspAsnProProAlaProAsnLysProThrPro	702
Qy	1151	TTTCAAGCCACCATCACAC---	1168
Db	703	ThrSerSerSerThrThrProProSerProAsnLeuProLeuGlnPheGlySerIleHis	722
Qy	1169	-----ATTGTAGAGGGAATCTTAGAGGTTAACATCATC---	1201
Db	723	SerProPheLeuSerAspGlyGlnLeuAsnTyrSerAlaLeuProProGlnAspPro	742
Qy	1202	---CAGATGACAGACGCTCCTCATGCGCGGTGCCATGGCTGAAAGCTCCCTAATAGACTTT	1255
Db	743	ThrAsnThrThrLeuSerLeuLeuProGluProLysPro-----	755
Qy	1259	GTGCTGACCTGCCAAGGAGGACATTCCCAAGGAGTTC---	1319
Db	756	-----ProThrGluValGlnSerProLeuMetAlaAspPro	767
Qy	1316	ACCTCGGAGATCCACCAACACACAGCTCTGCAGCCCTGGGATGTGGATGAGATGTCGTG	1375
Db	768	ThrCys-----ValGlyProAlaValSerPheSerSer---	778
Qy	1376	CTGACTGTGACAGAACCTTCAATGGGCTGGGACGTACTGTGTGAACCTCACCCCTGGG	1435
Db	779	LeuTyrProArgAspPhePheProAsnThrAlaSerPheLeuThrArgLeuArgLeuSer	798
Qy	1436	GATGACACAGCCTGGCTCTCAGAGCAC---	1480
Db	799	ProProThrProLeuProMetProLysAsnAsnCysLeuLeuThrAlaValAlaProSer	818
Qy	1487	GACCCAGCCTCCCTTTAAG---	1507
Db	819	LeuHisIleAsnProHisArgLeuThrThrSerLeuGlnGluValLeuProAspSerLeu	838
Qy	1508	ATGGCAAAACAGTGCCCTGATCTCCGTTGGCTGCTGGCCATATTTGTCACTGTGATCTCC	1567
Db	839	LeuSerAsnSerGluIleAspSerValGlyMetSerThrAspLeuLeuThrAlaLeuSer	858

US-10-039-272-1 (1-2661) x KGHUGH (1-525)

QY 1028 CCGCCACGCCACCA-----CCACCCAGACTTCAAAACCC 1063  
|||||  
Db 297 ProProProProAspGluArgAspHisHisGlyProProLeuProGlnGlyPro 316  
|||||  
QY 1064 ACCCTCTTTAGACCTGCTGTGTGACACCCCTGGAGCTGAGTAGGATTCCTGATGA 1123  
|||||  
Db 317 ProProLeuLeu-----ProMetSerCysSer----- 325  
|||||  
QY 1124 AACTGCCAGATTAAACAGATATGCCA----- 1149  
|||||  
Db 326 SerCysGlnHisAlaThrPheGlyThrAsnGlyAlaGlnArgHisSerHisAsnAsn 345  
|||||  
QY 1150 -----CTTTCAAGCCAC---CATCACAATTGTAGGGAAATCTTGTAGGTTAA 1197  
|||||  
Db 346 SerSerAspLeuHisProHisLeuHisHisSerHisGlnGlnHisProHisGly--His 364  
|||||  
QY 1198 CATCCAGATACACAGACGCTCTGTATGCGGTGCCATGCCTGAAAGCTCCCTAATAGACTT 1257  
|||||  
Db 365 HisProHisAlaHisHisProHisGlu----- 373  
|||||  
QY 1258 TGTGCTGACCTGCCAAGGGAGCATTCCTCCACGGAGGCTGTACCATCATTTCTGACCC 1317  
|||||  
Db 374 ---HisAspThrHisArgGlnHisProHisGly-----HisHis-----ProHis 387  
|||||  
QY 1318 CTGCGAGATCACCACGAGAACAGCTCTGCACGCCCTGTGATGTGATGAGATGTCTGCT 1377  
|||||  
Db 388 -----GlyHisHisProHisGlyHisHisProHisGly----- 398  
|||||  
QY 1378 GACTGTGAGACGAACCTTCAATGGTCTGGGACGCTACTGTGTGAACCTCACCTCGG 1437  
|||||  
Db 399 -----HisHisProHisGly----- 403  
|||||  
QY 1438 TGACACAAGCTGGCTCTCACGAGCACCCCTGATTTCTGTTCTGACAGAGACCCAGCCTC 1497  
|||||  
Db 404 ---HisHisProHisCysHisAsp----- 410  
|||||  
QY 1498 GCCTTTAAGATGGCAACACATGCCTGCTGCTGGCTGTGTCCTGATTCGCTGTCAT 1557  
|||||  
Db 411 ---PheGlnAspTyrGlyProCysAspProProHisAsnGlnGlyHisCysCysHis 429  
|||||

RESULT 15

QJ0533

genome polyprotein - Kennedyya yellow mosaic virus (strain Jarvis Bay)

N:Alternate names: RNA nucleotidyltransferase (RNA-directed); RNA replicase

N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)

C:Species: Kennedyya yellow mosaic virus

C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 19-Jan-2001

C:Accession: JQ0533

R: Ding, S.; Keese, P.; Gibbs, A.

J. Gen. Virol. 71, 925-931, 1990

A: Title: The nucleotide sequence of the genomic RNA of Kennedyya yellow mosaic tymovirus

A: Reference number: JQ0532; PMID: 90218040; PMID: 2324710

A: Accession: JQ0533

A: Molecule type: genomic RNA

A: Residues: 1-1874 <DIN>

A: Cross-references: GR:D00637; NID:G221969; PIDN: BAA00532.1; PID: G221971

C: Superfamily: eggplant mosaic virus RNA-directed RNA polymerase

C: Keywords: ATP; nucleotide binding; nucleotidyltransferase; P-loop; RNA biosynthesis; F

F: 1002-1009/Region: nucleotide-binding motif A (P-loop)

F: 1064-1069/Region: nucleotide-binding motif B

F: 1008/Binding site: ATP (Lys) #status predicted

Qy 1568 CTCTTG 1573  
|||  
Db 859 HisLen 860

Search completed: September 10, 2004, 15:51:13  
Job time : 108.5 secs



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2004, 01:04:08 ; Search time 10036 Seconds  
(without alignments)  
7917.822 Million cell updates/sec

Title: US-10-039-272-1  
Perfect score: 2661  
Sequence: 1 cggcacgagggccagagga.....aaaaaaaaaaaaaaaaaaaaa 2661

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estro:\*

7: em\_estro:\*

8: em\_estro:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_est3:\*

12: gb\_est4:\*

13: gb\_est5:\*

14: gb\_est6:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pln:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2550	95.8	2636	11	BC025297
2	1049.8	39.5	3615	11	AK044764
3	1046.6	39.3	2265	11	AK076347
4	1025	38.5	1201	13	BC397314

C	5	1014.6	38.1	1052	13	BC379978	BC379978
C	6	943.8	35.5	2431	11	AK079220	AK079220 Mus muscu
C	7	937	35.2	1003	13	BC423077	BC423077
C	8	926.6	34.8	1015	13	BC458448	BC458448
C	9	918.8	34.5	1139	13	BC406949	BC406949
C	10	914	34.3	1054	13	BC379979	BC379979
C	11	910.6	34.2	1201	13	BC336884	BC336884
C	12	906.4	34.1	1201	13	BC364871	BC364871
C	13	885.6	33.3	1012	13	BC458449	BC458449
C	14	868	32.6	1138	13	BC406948	BC406948
C	15	867.4	32.6	895	14	CF552020	CF552020 AGENCOURT
C	16	864.4	32.5	1002	13	BC423078	BC423078
C	17	862.2	32.4	927	9	AL542811	AL542811 AL542811
C	18	854	32.1	1096	13	BC381217	BC381217
C	19	852	32.0	1201	13	BC364870	BC364870
C	20	835.4	31.4	1048	9	AL575920	AL575920
C	21	831.6	31.3	972	13	BC396828	BC396828
C	22	823	30.9	923	9	AL542812	AL542812
C	23	816.6	30.7	891	13	BC452668	BC452668
C	24	815.8	30.7	911	12	BI521316	BI521316
C	25	799.8	30.1	972	13	BC396829	BC396829
C	26	796.8	29.9	858	12	BG742951	BG742951
C	27	788	29.6	1107	12	BM550298	BM550298
C	28	779	29.3	1089	12	BM547872	BM547872
C	29	776	29.2	835	12	BG576651	BG576651
C	30	775	29.1	808	12	BG742272	BG742272
C	31	775	29.1	962	12	BG676576	BG676576
C	32	758.2	28.5	810	9	AU139997	AU139997
C	33	750.4	28.2	1201	9	AL564958	AL564958
C	34	745	28.0	754	12	BG211992	BG211992
C	35	743.8	28.0	892	13	BU157659	BU157659
C	36	742.4	27.9	878	13	BC437190	BC437190
C	37	737.8	27.7	855	13	BU178305	BU178305
C	38	733.8	27.6	807	12	BG696053	BG696053
C	39	731.6	27.5	905	14	CD109017	CD109017
C	40	728.2	27.4	857	13	BU150501	BU150501
C	41	722.2	27.1	808	14	CD108927	CD108927
C	42	715.2	26.9	941	9	AL550419	AL550419
C	43	714.8	26.9	745	12	BG742920	BG742920
C	44	714.8	26.9	793	13	BU195030	BU195030
C	45	714.8	26.9	891	13	BC412287	BC412287

## ALIGNMENTS

RESULT 1	BC025297	Homo sapiens, glycoprotein (transmembrane)	nmb, clone	linear	HTC 08-MAR-2002
LOCUS	BC025297	IMAGE:4877773, mRNA.			
DEFINITION	BC025297	BC025297.1	GI:19264140		
ACCESSION	BC025297	HTC.			
VERSION	BC025297.1	Homo sapiens (human)			
KEYWORDS	HTC.	Homo sapiens			
SOURCE	HTC.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
ORGANISM	HTC.	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1	(bases 1 to 2636)			
AUTHORS		Strausberg,R.			
TITLE		Direct Submission			
JOURNAL		Submitted (05-MAR-2002) National Institutes of Health, Mammalian			
REMARK		Gene Collection (MGC), Cancer Genomics Office, National Cancer			
COMMENT		Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,			
		USA			
		NIH-MGC Project URL: http://mgc.nci.nih.gov			
		Contact: MGC help desk			
		Email: cgabbs@mail.nih.gov			
		Tissue Procurement: ATCC/DCTD/BTP			
		cDNA Library Preparation: Rubin Laboratory			
		DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LML)			
		BC Cancer Agency, Vancouver, BC, Canada			

info@bcsc.bc.ca  
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
Leticia Hsiao, Martin Krywinski, Reta Kutsche, Oliver Lee, Soo  
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
Nees, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline  
Scheln, Duane Smaluis, Michael Smith, Lorraine Spence, Jeff Stott,  
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,  
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 42 Row: 1 Column: 5  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 4505404  
This clone has the following problem: frame shifted.

FEATURES  
Location/Qualifiers  
1..2636  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="LocusID:10457"  
/db\_xref="taxon:9606"  
/clone="IMAGE:487773"  
/tissue\_type="Skin, melanotic melanoma, high MDR."  
/clone\_lib="NIH\_MGC\_49"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"

ORIGIN  
Query Match 95.8%; Score 2550; DB 11; Length 2636;  
Best Local Similarity 99.1%; Pred. No. 0;  
Matches 2595; Conservative 0; Mismatches 20; Indels 3; Gaps 3;

QY	28	AACCTTGGTGGCTGGCTGGTGGATTCAGATGGAATTCAGATGGAATGCTCTACTATTTTCCTGGGAT	87
DB	20	ACCTTGGTGGCTGGCTGGTGGATTCAGATGGAATGCTCTACTATTTTCCTGGGAT	79
QY	89	TTCTGCTCCTGGCTGGAGATTCGCACTTGATGCGCGCAAGATTCATGATGCTGG	147
DB	80	TTCTGCTCCTGGCTGGAGATTCGCACTTGATGCGCGCAAGATTCATGATGCTGG	139
QY	148	GCAATGAAGACCTTCTGCTTACATGAGGAGCACAATCAATTAATGGCTGCTCTG	207
DB	140	GCAATGAAGACCTTCTGCTTACATGAGGAGCACAATCAATTAATGGCTGCTCTG	199
QY	208	ATGAAATGACTTGAATGAAATCTACCCAGTGTGGAAGCGGGGAGACATGAGTGA	267
DB	200	ATGAAATGACTTGAATGAAATCTACCCAGTGTGGAAGCGGGGAGACATGAGTGA	259
QY	268	AAACTCTGGAGGAGCGGTGTCAGCGGCTGCTGACCTGACTCACCAGCCCTCG	327
DB	260	AAACTCTGGAGGAGCGGTGTCAGCGGCTGCTGACCTGACTCACCAGCCCTCG	319
QY	328	TGGGCTCAATATAAATTAATTTGGGTGAACCTGATATTCCTAGATGCCAAAGAGATG	387
DB	320	TGGGCTCAATATAAATTAATTTGGGTGAACCTGATATTCCTAGATGCCAAAGAGATG	379
QY	388	CCAATGGCAACATAGTCTATGAGAGAACTGCAAGAAATGAGGTGTTATCTGCTGATC	447
DB	380	CCAATGGCAACATAGTCTATGAGAGAACTGCAAGAAATGAGGTGTTATCTGCTGATC	439
QY	448	CATATGTTTCAACTGACAGCATGTCAGAGCAGATGACGGGAAATGGCACCGGCC	507
DB	440	CGTATGTTTCAACTGACAGCATGTCAGAGCAGATGACGGGAAATGGCACCGGCC	499
QY	508	AAAGCCATCATACGCTTCCCTGATGGAAACCTTTTCTCACCACCCCGGATGGAGAA	567
DB	500	AAAGCCATCATACGCTTCCCTGATGGAAACCTTTTCTCACCACCCCGGATGGAGAA	559
QY	568	GATGGAATTCATCTACGCTTCCACACATCTGGTGTGATTTTCCAGAAATTTGGGAGAT	627
DB	560	GATGGAATTCATCTACGCTTCCACACATCTGGTGTGATTTTCCAGAAATTTGGGAGAT	619

QY	628	GTTCAAGTGAAGATTTTCTGTGAACACACACCAATGTGACACTTGGGCTTCAACTCATGGAAG	687
DB	620	GTTCAAGTGAAGATTTTCTGTGAACACACACCAATGTGACACTTGGGCTTCAACTCATGGAAG	679
QY	688	TGACTGTCTCAGAGAATGAGCGGCATATGTTCCCATCGCAAGTGAAGATGATGTT	747
DB	680	TGACTGTCTCAGAGAATGAGCGGCATATGTTCCCATCGCAAGTGAAGATGATGTT	739
QY	748	ACGTGTGTAAACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAAAGAACCATCGAAAT	807
DB	740	ACGTGTGTAAACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAAAGAACCATCGAAAT	799
QY	808	CATCCGACGAAACCTTCC-CAAGATCTCCCATTTATGTTTGTGATGCTCTGATTCATGATC	866
DB	800	CATCCGACGAAACCTTCC-CAAGATCTCCCATTTATGTTTGTGATGCTCTGATTCATGATC	859
QY	867	CTAGCCACTTCTCTCAATTTTCTTACCATTAATCTCAAGTGGAGCTTCGGGGATATATCTG	926
DB	860	CTAGCCACTTCTCTCAATTTTCTTACCATTAATCTCAAGTGGAGCTTCGGGGATATATCTG	919
QY	927	GCCTGTTTGTTCACCAATCATATCTGTGAATCACACGTATGTGCTCAATGGAACCTTCA	986
DB	920	GCCTGTTTGTTCACCAATCATATCTGTGAATCACACGTATGTGCTCAATGGAACCTTCA	979
QY	987	GCCTTAACTCACTGTGAAAGCTGCAGCACCAAGACCTTGTCCGCGACCGCCACCAAC	1046
DB	980	GCCTTAACTCACTGTGAAAGCTGCAGCACCAAGACCTTGTCCGCGACCGCCACCAAC	1039
QY	1047	CGAGACCTTCAAAACCCACCTTCTTTAGGACCTGCTGTGACACACCCCTGGAGCTGA	1106
DB	1040	CGAGACCTTCAAAACCCACCTTCTTTAGGACCTGCTGTGACACACCCCTGGAGCTGA	1099
QY	1107	GTAGGATTCCTGATGAAACTGCCAGATTAACAGATATGGCCACTTTTCAAGCCACCATCA	1166
DB	1100	GTAGGATTCCTGATGAAACTGCCAGATTAACAGATATGGCCACTTTTCAAGCCACCATCA	1159
QY	1167	CAATTTAGAGGGAATCTTTAGAGGTTAAACATCATCCAGATGACAGCTTCCTGATGCTGG	1226
DB	1160	CAATTTAGAGGGAATCTTTAGAGGTTAAACATCATCCAGATGACAGCTTCCTGATGCTGG	1219
QY	1227	TGCCATGGCTGAAAGCTCCCTTAATAGACTTTTGTGTCGACCTGCCAAGGGAGCATTCCTCA	1286
DB	1220	TGCCATGGCTGAAAGCTCCCTTAATAGACTTTTGTGTCGACCTGCCAAGGGAGCATTCCTCA	1279
QY	1287	CGAGGCTGTACCATCATTTCTGACCCACCTGCGAGATCACCCAGAACACAGTCTGCA	1346
DB	1280	CGAGGCTGTACCATCATTTCTGACCCACCTGCGAGATCACCCAGAACACAGTCTGCA	1339
QY	1347	GCCCTGTGATGTGAGATGATGTTGCTGCTGATGTGAGAGCAACCTTCAATGGGCTG	1406
DB	1340	GCCCTGTGATGTGAGATGATGTTGCTGCTGATGTGAGAGCAACCTTCAATGGGCTG	1399
QY	1407	GGAAGTGTGTGAGACTCAGCTCAGCTGGGGATGACACAGCTTGGCTCTCAGGAGACCC	1466
DB	1400	GGAAGTGTGTGAGACTCAGCTCAGCTGGGGATGACACAGCTTGGCTCTCAGGAGACCC	1459
QY	1467	TGATTTCTCTTCTGACAGAGACCCAGCTTGGCTTAAAGGATGGCAACAGTCCCTGA	1526
DB	1460	TGATTTCTCTTCTGACAGAGACCCAGCTTGGCTTAAAGGATGGCAACAGTCCCTGA	1519
QY	1527	TCTCCGTTGGCTGCTGGCCATTTTGTGATGATCTCCCTCTTGGTGTACAAAAAC	1586
DB	1520	TCTCCGTTGGCTGCTGGCCATTTTGTGATGATCTCCCTCTTGGTGTACAAAAAC	1579
QY	1587	ACAAGAAATACACCAATAGAAATAGTCTGGGATGTGTCAGAGCAAGAAAGGCTGTA	1646
DB	1580	ACAAGAAATACACCAATAGAAATAGTCTGGGATGTGTCAGAGCAAGAAAGGCTGTA	1639
QY	1647	GTCTCTTTTCTCAACCGTGAAGAGCGGTGTTTCCCGGAAACCCAGGAAAGGATCCGC	1706
DB	1640	GTCTCTTTTCTCAACCGTGAAGAGCGGTGTTTCCCGGAAACCCAGGAAAGGATCCGC	1699
QY	1707	TACTCAAAACCAAGAAATTTAAAGGAGTTTCTTAATTTTCGACCTTCTTCTGAAGCTCA	1766



```
Db 1700 TACTCAAAACCAAGAAATTTAAGAGATTCTTAATTTTCGACCTGTTCTTGAGACTCA 1759
Qy 1767 CTTTTCAGTGCCATTGATGTGAGATGTGCTGGAGTGGCTATTAACTTTTTTTCCTAAAG 1826
Db 1760 CTTTTCAGTGCCATTGATGTGAGATGTGCTGGAGTGGCTATTAACTTTTTTTCCTAAAG 1819
Qy 1827 ATTATTGTTAAATAGATATTGGTTGGGGAAGTTGAATTTTTTATAGTTAAATGTCA 1886
Db 1820 ATTATTGTTAAATAGATATTGGTTGGGGAAGTTGAATTTTTTATAGTTAAATGTCA 1879
Qy 1887 TTTTAGAGATGGGAGAGGATTTACTGACGAGCTTCAGCCATGTTGCTAAAGCTGAT 1946
Db 1880 TTTTAGAGATGGGAGAGGATTTACTGACGAGCTTCAGCCATGTTGCTAAAGCTGAT 1939
Qy 1947 AAAACAACCTTAGCAAGGCTCTTTTCATTATTTTATGTTTCACTTATTAAGCTTTAG 2006
Db 1940 AAAACAACCTTAGCAAGGCTCTTTTCATTATTTTATGTTTCACTTATTAAGCTTTAG 1999
Qy 2007 GTAACTAGTAGGATAGAAACACTGTGTCGAGAGCTTAAGGAGAGCTACTATTGATTA 2066
Db 2000 GTAACTAGTAGGATAGAAACACTGTGTCGAGAGCTTAAGGAGAGCTACTATTGATTA 2059
Qy 2067 GAGCTTAACCAAGGTTAACTCAAGAGAGGCGGATACTTTCAGCTTTCATGTAAGCTG 2126
Db 2060 GAGCTTAACCAAGGTTAACTCAAGAGAGGCGGATACTTTCAGCTTTCATGTAAGCTG 2119
Qy 2127 TATGCATTAAGCCCAATGTAGTCCAGTCTTCAAGATCATGTTCCAAAGCTAACTGCCA 2186
Db 2120 TATGCATTAAGCCCAATGTAGTCCAGTCTTCAAGATCATGTTCCAAAGCTAACTGCCA 2179
Qy 2187 CTTCAATACACACTCATGAACCTCTGTGAGAACATACAGGCCCAAGCTGTGGTATGA 2246
Db 2180 CTTCAATACACACTCATGAACCTCTGTGAGAACATACAGGCCCAAGCTGTGGTATGA 2239
Qy 2247 TGTGCACACTTGTAGACTCAGAAAAAATACTCTCATAAATGGGTGGGAGTATTTTG 2306
Db 2240 TGTGCACACTTGTAGACTCAGAAAAAATACTCTCATAAATGGGTGGGAGTATTTTG 2299
Qy 2307 GTGCACACTTGTAGTCTGCTGAGTGAAGGAATGATATTCATATTCATTTATTCCTCA 2366
Db 2300 GTGCACACTTGTAGTCTGCTGAGTGAAGGAATGATATTCATATTCATTTATTCCTCA 2359
Qy 2367 TGGACATTAGTATGAGTCTTTTATATACAGGATGATGCTGACACACTTTGTGTA 2426
Db 2360 TGGACATTAGTATGAGTCTTTTATATACAGGATGATGCTGACACACTTTGTGTA 2419
Qy 2427 TATTTCCAAATTTTGTATAGTCTGCACATATTTGAAATCAAAATATTAAGACTTTCC 2486
Db 2420 TATTTCCAAATTTTGTATAGTCTGCACATATTTGAAATC-ATATATTAAGACTTTCC 2478
Qy 2487 AAAAAATTTGCTCCCTGTTTTCATGCAACTGATCAGTAAGGATTTCCCTCTGTTTG 2546
Db 2479 AAAGATGAGGTCCTGTTTTCATGGCAACTGATCAGTAAGGATTTCCCTCTGTTTG 2538
Qy 2547 GAACATAAACCACTTACTATATGTTAGACAGACATTTTTCCTTCTCCTCTGAAAAA 2606
Db 2539 TAACTAAACCACTTACTATATGTTAGACATGATCTTTTCTCTCCTCTGAAAAA 2598
Qy 2607 -AAATGAGGGAAGAGACAAAAAATAAAAAAATAAAAAA 2643
Db 2599 TAAAGTGTGGAAGAGACAAAAAATAAAAAAATAAAAAA 2636
```

RESULT 2  
AK044764

LOCUS  
AK044764

DEFINITION  
Mus musculus adult retina cDNA, RIKEN full-length enriched library,  
clone:A930040L12 product:glycoprotein (transmembrane) mmb, full  
insert sequence.

ACCESSION  
AK044764

VERSION  
AK044764.1

KEYWORDS  
HTC; CAP trapper.

SOURCE  
ORGANISM

REFERENCE  
AUTHORS

TITLE  
JOURNAL

MEDLINE  
PUBMED

REFERENCE  
AUTHORS

TITLE  
JOURNAL

MEDLINE  
PUBMED

REFERENCE  
AUTHORS

TITLE  
JOURNAL

MEDLINE  
PUBMED

REFERENCE  
AUTHORS

TITLE  
JOURNAL

MEDLINE  
PUBMED

REFERENCE  
AUTHORS

TITLE  
JOURNAL

MEDLINE  
PUBMED

REFERENCE  
AUTHORS

TITLE  
JOURNAL

MEDLINE  
PUBMED

REFERENCE  
AUTHORS

TITLE  
JOURNAL

MEDLINE  
PUBMED

REFERENCE  
AUTHORS

TITLE  
JOURNAL

MEDLINE  
PUBMED

REFERENCE  
AUTHORS

TITLE  
JOURNAL

MEDLINE  
PUBMED

REFERENCE  
AUTHORS

COMMENT

Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)  
99279253  
10349636  
2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
20499374  
11042159  
3 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kitsumi, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861  
4 The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)  
5 The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 3615)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayaishi, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akai, S.,  
Takeda, Y., Tanaka, T., Tomaru, A., Taya, T., Yasunishi, A.,  
Yamada, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Retina RNA was provided by Dr. Stefano Gustincich (Department of  
Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA  
02115, USA) whose assistance is gratefully acknowledged. Please  
visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.ssc.riken.go.jp/. Location/Qualifiers		510	AGCCATCATTAAGCGTCTTCCCTGATGGGAAACCTTTCTCTACCAACCCCGGATGAGAGA	569
FEATURES	source	QY		
	1..3615	Db		609
	/organism="Mus musculus"	QY		
	/mol_type="mRNA"	Db		629
	/strain="C57BL/6J"	QY		
	/db_xref="FANTOM_DB:A930040L12"	Db		669
	/db_xref="MGI:2410100"	QY		
	/clone="A930040L12"	Db		689
	/tissue_type="retina"	QY		
	/clone_lib="RIKEN full-length enriched mouse cDNA library"	Db		729
	/dev_stage="adult"	QY		
	100..1824	Db		749
CDS	/note="unnamed protein product; glycoprotein (transmembrane) nmb (MGI MGI:1934765, GB NM_053110, evidence: BLASTN, 99%, match=2272) putative"	QY		
	/codon_start=1	Db		809
	/protein_id="BAC32074.1"	QY		
	/db_xref="GI:26336783"	Db		849
	/translation="MESLCGVILGFLLLAAGLPLOAKRFRDVLGHEQVYNHMRHNQL RGWSDENWEHLVPMVRGDRWKDGLVNTAGADGDWEDGTSRQHLRFPDRR PRCQEDANGIVVEKNCNDLGLTSLHVNMTAGADGDWEDGTSRQHLRFPDRR FFPRPHGKKWSFVVFHLLGQYFQKLRCSARVSNINTVLTAGPQVMEVTVFRYGR AYIPISKVDVVTIDQIPVFTMSQKNDRLNLSDEIFLRLDPIVEDVLLHPSHFLND SAISYKWNFGDNTGLFVNNHNTLNHYVINGTLFNLITVQTAVPFCPPFPSTFPFP STPSPSPSPPTLTPSPSLMPTGYKSMELSDISNENCRINRYGYFRATITVEGIL EVSIMQIADVPMTPOPANLMDFTVCKGAPMEACTIISDPTQIQAQNRVCSFVAV DGLCLLSVRRAVPGSTYCNVFTLGDASLALSTLSIPGKDPSPMRVNGVLISI GCLAVLNTVITLLYKKHKAYKPIGNCPNRTVEGKLSVLLSHAKAPFRFGQDEKDL LQDKPRTL"	QY		928
	3600..3605	Db		969
	/note="putative"	QY		
	3615	Db		1029
	/note="putative"	QY		
	3615	Db		1037
ORIGIN		QY		
		Db		1038
		QY		
		Db		1060
		QY		
		Db		1090
		QY		
		Db		1149
		QY		
		Db		1120
		QY		
		Db		1209
		QY		
		Db		1180
		QY		
		Db		1269
		QY		
		Db		1240
		QY		
		Db		1329
		QY		
		Db		1300
		QY		
		Db		1389
		QY		
		Db		1360
		QY		
		Db		1449
		QY		
		Db		1420
		QY		
		Db		1509
		QY		
		Db		1480
		QY		
		Db		1569
		QY		
		Db		1540
		QY		
		Db		1629

QY 1541 TTGGCCATATTGTCTACTGTGATCTCCCTCTGGTGTACAAAAACACAAAGGAATACAC 1600  
 |||||  
 Db 1630 CTGGCTGTGCTGTACCAATGTTACCATCTTCTGTGTACAAAAACACAGGCGTACAG 1689  
 |||||  
 QY 1601 CCAATAGAAAATAGTCTCGGTAATGTGTGTCAGAAACAAAGGCTGAGTCTCTTCTCAAC 1660  
 |||||  
 Db 1690 CCAATAGGAAATGCTCCCAAGAACACGCTCGAGGGCAAAAGGCTGAGTCTCTCTCAGC 1749  
 |||||  
 QY 1661 CCGTCAAAAGCCGTGTTCTTCCCGGGAACACAGGAAGGATCCGCTACTCAAAACCAA 1720  
 |||||  
 Db 1750 CACGCGAAAGCCCGTCTTCGAGAGACCGAGGAGAGATCCATTCGTCAGGACCAAG 1809  
 |||||  
 QY 1721 GAATTTAAAGGAGTTTCTTAAATTTTCGACTGTTCTTGAAGCTCACTTTTCAGTGCAT 1780  
 |||||  
 Db 1810 CCAAGGACACTTAAGTCTTTTGGCCTTCCTCTGACCAGGACCACTCTTCTTGIG-CAT 1868  
 |||||  
 QY 1781 TGATGTGAGATGCTGTGA----GTGGCTATTAACCTTTTCTCTAAGATTAATTGTA 1836  
 |||||  
 Db 1869 GTATGTGAGCTGTGCAGAAATGATGTGGCTGGGAAGTGTGTTCTCTAAGGATTAATTGTA 1928  
 |||||  
 QY 1837 AATGATATTGTGGTTTGGGGAAGTTGAATTTTATAGGTTAAA 1881  
 |||||  
 Db 1929 AATGATATTGTGGCTTGGGAGTGTGTTAAATAGCATTTTGA 1973  
 |||||

RESULT 3  
 AK076347  
 LOCUS  
 DEFINITION Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched library, clone:4732452017 product:glycoprotein (transmembrane) nmb, full insert sequence.  
 AK076347  
 VERSION AK076347.1 GI:26345309  
 KEYWORDS HTC; CAP trapper  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1  
 Carninci, P., and Hayashizaki, Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636  
 2  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20499374  
 PUBMED 11042159  
 3  
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishize, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE 20530913  
 PUBMED 11076861  
 4  
 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
 TITLE Functional annotation of a full-length mouse cDNA collection  
 JOURNAL Nature 409, 685-690 (2001)  
 5  
 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 JOURNAL Nature 420, 563-573 (2002)  
 REFERENCE 6 (bases 1 to 2265)  
 AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
 COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Please visit our web site for further details.  
 URL: http://genome.gsc.riken.go.jp/  
 URL: http://fantom.gsc.riken.go.jp/  
 Location/Qualifiers  
 1. 2265  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="FANTOM DB:4732452017"  
 /db\_xref="MGI:2350753"  
 /db\_xref="taxon:10090"  
 /clone="4732452017"  
 /tissue\_type="skin"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="10 days neonate"  
 75. 1799  
 /note="unnamed protein product; glycoprotein (transmembrane) nmb (MGI:1934765, GB|NM\_053110, evidence: BLASTN, 99%, match=2272)"  
 putative  
 /codon\_start=1  
 /protein\_id="BAC36306.1"  
 /db\_xref="GI:26345310"  
 /translation="MESLGVGLFLLAAGLEPQAAKRFVDLGHQYQPDHMEHNLKRWSSDENWDEHLVPMWRGGRWKSWEGRVQAVLTSDSALVGSNITFYVNLVFPCKQEDANGNIYKXNNDLGLTSDLHYAWTAGADGDWEDGT'SRSHLRFPPRRPFRPHGKXWVYVFTLQGYFQKLGCSARVSINTVNLTAGPOVMTVFRYGRATVPSKVDVYVTDQIPVFTMSOKNDRNLSDEIFLRDLPIVEVLVHDPHFLLND SALSXWNGDNTGLFVSNHNTLHYVNLNGTFLNLTQVTAVPGCPSPSPSTPPPP STPPSPPLTSLTSPSLMPTGYKSMELSDISNENCRINRYFRATITVEGIL EYSLMIADVPMTPOPANLMDFTVTCGATPMEACTTISDPTQIAQNRVGVAV DGLCLSVRAFNGSGTCVNFITGLDASLALTSLISIPGKDPDPLRAVNGVLISI GCLAVLTMTVITLLYKKHAYKPIGNCPNTVKGKGLSVLLSHAKAPFFRGQDKPL LQDKPRTL"  
 2250. 2255  
 /note="putative"  
 2265  
 /note="putative"  
 polyA\_signal  
 2250. 2255  
 polyA\_site  
 2265  
 /note="putative"  
 ORIGIN  
 Query Match 39.3%; Score 1046.6; DB 11; Length 2265;  
 Best Local Similarity 73.9%; Pred. No. 1.9e-159;  
 Matches 1407; Conservative 0; Mismatches 444; Indels 54; Gaps 4;

Qy 30 CTTGGTGGCTGGTGGATTCAGCATGGAATGTCTCTACTATTTCCTGGGATTT 89  
Db 45 CTTGACTGGTTCAGGGCGCTGGAGTCAGCATGGAAGTCTCTCGGGGTCTCGGATTT 104  
Qy 90 CTGCTCTGGTTCGAAGATTGCCACTTGTATGCCGCCCAAAAGATTTCATGATGTCTGGGC 149  
Db 105 CTGCTCTGGTTCGAGAGACTGCTCTCCAGCTGCGCAAGCGATTTCGTGATGTCTGGGC 164  
Qy 150 AATGAAAGACCTCTCTGCTTACATGAGGGAGCACAAATCAATTAATTAATGCTGTCTTCTGAT 209  
Db 165 CATGAAACAGTATCCCGATCAATGAGAGAGACAAACAAATTAACCTGGTGTCTTCTCGAT 224  
Qy 210 GAAATGATGGAATGAATAACTCTACCACTGTGGAGCGGGAGACATGAGTGGAAA 269  
Db 225 GAAATGAAATGGGATGAACACCTGTATCCAGTGTGGAGGGGAGACGCGCAGGTGGAG 284  
Qy 270 AACTCTCTGGAAGGAGGCGGTGTGAGGCGGTCTCTGACCACTGACTCACACCGCTCTGGT 329  
Db 285 GACTCTCTGGAGAGGCGGTGTGAGGCGGTCTCTGACCACTGACTCACACCGCTCTGGT 344  
Qy 330 GGCTCAAAATATAACATTTGGCGTGAACTCTGATATTTCCCTAGATGCGCAAAAGGAGATGCC 389  
Db 345 GGTTCCAATATACATTTTGTGTGAACCTGTGTTCCTCCAGATGCGCAAGGAAATGCT 404  
Qy 390 AATGCAACATAGTCTATGAGNAGAACTGCGAATGAGGCTGTTTATCTGCTGATCCA 449  
Db 405 AATGCAATATGCTATGAGAAGAACTGCGAATGAGTATGAGGACTGACATCTGACCTG 464  
Qy 450 TATGTTTAACTGACAGCATGCTCAGAGGACAGTGACGGGGAATAATGGCACCGGCCAA 509  
Db 465 CATGCTCAACTGACATGCGAGGGCAGATGATGCTGAGTGGAAAGATGGCACCGCGCA 524  
Qy 510 AGCCATCATAACTCTTCCCTGATGGGAACTCTTTTCTCAGACACCGCGGATGGAGAGA 569  
Db 525 AGCCAGCATCTCAGGTTCCCGACAGAGGCGCTTCCCTCGCCCGCATGGATGGAGAAA 584  
Qy 570 TGGAAATTCATCTAGTCTTCCACACACTGCTGCTAGTATTTCCAGAAATGGGACGATG 629  
Db 585 TGGAGCTTGTCTAGTCTTTTCCACACACTGCTGCTAGTATTTCCAGAAATGGGACGATG 644  
Qy 630 TCAGTGAGAGTCTTCTGTAACACAGCAATGTGACACTTGGGCTCAACTCATGGAAGTG 689  
Db 645 TCAGCAGGGTCTTCTATAACACAGCTCACTTGGACAGTGGCCCTCAGGTCATGGAAGTG 704  
Qy 690 ACTGCTCAGAGACATGGAAGGCGATATGTTCCCATCGCACAGTGAAGATGTGTAC 749  
Db 705 ACTGCTTTCGAAGATACGCGCGGCGATACATTTCCCATCTCGAAGGTGAAGATGTGTAT 764  
Qy 750 GTGGTAACAGATCAGATTCCTGTTGTTGTGACTATGTTCCAGAGAACGATCGAAATTC 809  
Db 765 GTGATAACAGATCAGATTCCTGTTGTTGTGACTATGTTCCAGAGAACGATGACAGAACTTG 824  
Qy 810 TCCGACGAAACCTTCC - CAAAGATCTCCCATATGTTGATGTCCTGATTCATGATCCT 868  
Db 825 TCTGATGAGATCTTCTCAGAGACCTCCCATCGTCTTCGATGTCCTCATTCATGATCCC 884  
Qy 869 AGCCACTTCTCAATTTACTACATTAACATACAGTGGAGCTTCGGGATATATCTGGC 928  
Db 885 AGCCACTTCTCAACGACTCTGCCATTTCTCAAGTGGAACTTTGGGGCAACACTTGGC 944  
Qy 929 CTGTTTGTTCACCAATCATACTGTGAATCACAGCTATGCTCAATGGAACCTTCAGC 988  
Db 945 CTGTTTGTCTCCAAACATACACTTTGAATCACACTATGTTGCTCAATGGAACCTTCAAC 1004  
Qy 989 CTTAACTCACTGGAAGCTCGAGCACAGGACCTTGTCCGCCACGCGC----- 1037  
Db 1005 CTTAACTCACTGGAAGCTCGAGCACAGGACCTTGTCCGCCACGCGC----- 1064  
Qy 1038 -----CACCACACCGACCTTCAAAA 1060  
Db 1065 CCGCTCCACTTCACTCCGCGCTCAGCTCCGCGCTCAGCTCTGCGCCATATCAACA 1124  
Qy 1061 CCCACCCCTTCTTTAGGACCTGCTGGTGAACAAACCCCTGGAGCTGAGTAGGATTCCTGAT 1120

Db 1125 CCTAGCCCTCTTTAATGCTCTACTGGTTACAAATCCATGAGCTGAGTGAATTCCTCAAT 1184  
Qy 1121 GAAATCTCCAGATTAACAGATATGGCCACTTTCAAGCCACCATCAAAATTTGTAGAGGA 1180  
Db 1185 GAAATCTCCAGATTAACAGATATGGCTACTTTCAGAGCCACCATCAAAATTTGTAGAGGG 1244  
Qy 1181 ATCTTAGAGTTAAATCATCATCCAGATGACAGACGTCTCTGATGCGGGTCCCATGGCCTGAA 1240  
Db 1245 ATCTGGAAGTCAGCATCATGAGATAGCAGATGTCCTCCATGCGCCACACCGCAGCTGCC 1304  
Qy 1241 AGCTCCCTAATAGACTTTGTGTCACCTGCCAAGGGAGCATTCCTCCAGAGGTCTGTATCC 1300  
Db 1305 AACTCTCTGATGAGCTTCACTGTGACCTGCAAGGGGCGCCATGCGCATGGAAGCTGTAGC 1364  
Qy 1301 ATCAATTTCTGACCCCACTGCGAGATCACCCAGAAACAGCTCTGAGCGCTGTGTGATGTG 1360  
Db 1365 ATCAATCTCCGACCCCACTGCGAGATCGCCAGAACCGGGTCTGAGCGCTGTGGCTGTG 1424  
Qy 1361 GATGAGATGTGTCTGCTGACTGTGAGAGAACCTTCAATGAGGTCTGGGACGTACTGTGTG 1420  
Db 1425 GATGGCTGTGCTGTCTGTGAGAAAGACCTTCAATGAGGTCTGGCACCTACTGTGTG 1484  
Qy 1421 AACCTCACCTGGGGATGACAAAGCCTGCTCTCAGAGCACCTGATCTCCCTTGGCTGC 1480  
Db 1485 AATTTCACTCTGGGAGATGATCAAGCCTGGCCCTCACAGCACCTGATCTCTATCCCT 1544  
Qy 1481 GACAGAGACCCAGCTCGCTTTTAAGGATGGCAACAGTGGCCCTGATCTCCCTTGGCTGC 1540  
Db 1545 GCGAAAGACCCAGACTCCCTCTGAGAGAGTGAATGTGCTGATCTCCATCGGCTGC 1604  
Qy 1541 TTGGCCATATTTGTCACTGTGATCTCTCTTGGTGTACAAAAACAGAGAAATACAAAC 1600  
Db 1605 CTGGCTGTGCTGTCAACCATGTTTACCATCTTGTGTACAAAAACAGAGCGTACAAAG 1664  
Qy 1601 CCAATAGAAATAGTCTCTGGGAAATGTGTGACAGAAAGGCTGAGTGTCTTCTCAAC 1660  
Db 1665 CCAATAGGAAATCTGCCCCAGGAAACAGGTCAAGGCAAGGGCTGAGTGTCTCTCTCAGT 1724  
Qy 1661 CGTGCAAAAGCCGCTCTCTCCCGGAAACAGGAAAGGATCCCGTCACTAAAAACCAA 1720  
Db 1725 CACGCGAAAGCCCGCTCTCTCCAGAGACACAGAGAGAGGATCCATGCTCCAGGACAAAG 1784  
Qy 1721 GAATTTAAAGGAGTTTCTTAAATTTGACCTTGTGTTTGAAGCTCACTTTTCTAGTCCAT 1780  
Db 1785 CCAAGGACACTCTAAGTCTTTGGCCCTTCCCTGACAGGAAACCCACTCTTCTGTG - CAT 1843  
Qy 1781 TGATGTGAGATGTGCTGGA ---GTGGCTATTAACTTTTCTTAAAGATTATTGTTA 1836  
Db 1844 GTATGTGAGCTGTGAGAAAGTATGTGGCTGGGAACTGTGTTCTCTAAGGATTATTGTAA 1903  
Qy 1837 AATAGATATTGTGGTGTGGGAAAGTTGAAATTTTATAGGTTAAA 1881  
Db 1904 AATGATATCTGGCTGTAGGAGTGTGGTTAAATAGCAITTTTGA 1948

RESULT 4  
BX397314/c BX397314  
LOCUS BX397314 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
DEFINITION clone CS0D1033YF05 3-PRIME, mRNA sequence.  
ACCESSION BX397314  
VERSION BX397314.1 GI:30629122  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1201)  
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope

Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 4395.f For  
more information about this cluster, see

cgi-bin/cluster.cgi?seq=CSODI033CC03NP1&cluster=4396.f. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
<http://fulllength.invitrogen.com> URL :  
Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS9D1033CC03NP1.

## FEATURES

```

source
1. 1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSDDI033YF05"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

## ORIGIN

Query Match	38.5%	Score 1025;	DB 13;	Length 1201;
Best Local Similarity	95.9%	Pred. No. 7.9e-156;		
Matches 1072; Conservative	16;	Mismatches 26;	Indels 4;	Gaps 4;

QY	1499	CC	TTTAGGATGGCAACAGTGCCTGATCTCCGTTGGCTTGGCCATATTTGTCACT	1558
Db	1115	CC	TTTARGATGCAAAACASCGCCCTGATCTCTG-TGGCTGCTTGGCCATATTTGTCACT	1057
QY	1559	GT	GATCTCCCTCTTGGTGTCAAAAACACAGAGGAATACACCCCAATAGAAAAATGTCCT	1618
Db	1056	GT	ATCTCCCTC-TGGTGTCAAAAACACAAAGAAWACAACCAATAGAAAAATGTCCT	998
QY	1619	GG	AAATGTGTGTGAAGCAAAAGGCGCTGAGTGTCTTTCTC-AACCGTGCAAAAGCCGTT	1677
Db	997	GG	AAATGTGTTCAGAACAAAGGCGCTGAGTGTCTGTCTCAACCGTGCAAAAGCCGTT	938
QY	1678	CT	TCCCGGGAAACGAGAAAAAGATCCGCTACTCAAAAAACAAGAAATTTAAAGAGTTTC	1737
Db	937	CT	TCCCGGGAAACGAGAAAAAGATCCGCTACTCAAAAAACAAGAAATTTAAAGAGTTTC	878
QY	1738	TT	AAATTCGACCTGTCTTCTGAAGCTCACTTTTCAGTGCCATTGTGAGATGTGCTG	1797
Db	877	TT	AAATTCGACCTGTCTTCTGAAGCTCACTTTTCAGTGCCATTGTGAGATGTGCTG	818
QY	1798	GAG	TGGCTATTACCTTTTTTTCCTAAGATTTATGTTAAATAGATTTGTGTTTGGGG	1857
Db	817	GAG	TGGCTATTACCTTTTTTTCCTAAGATTTATGTTAAATAGATTTGTGTTTGGGG	758
QY	1858	AA	GTTGAAATTTTATAGGTTAAATGTCATTTTAGAGATGGGAGAGGATATATCTGCA	1917
Db	757	AA	GTTGAAATTTTATAGGTTAAATGTCATTTTAGAGATGGGAGAGGATATATCTGCA	698
QY	1918	GG	CAGCTTCAGCCATGTTGTGAACTGATAAAGCAACTTAGCAAGGCTTCTTTTCATTA	1977
Db	697	GG	CAGCTTCAGCCATGTTGTGAACTGATAAAGCAACTTAGCAAGGCTTCTTTTCATTA	638
QY	1978	TT	TTTTATGTTTCACATTATAAGCTTTAGTAACTAGTAGTAGGATAGAAAACACTGTGTCCCG	2037
Db	637	TT	TTTTATGTTTCACATTATAAGCTTTAGTAACTAGTAGTAGGATAGAAAACACTGTGTCCCG	578
QY	2038	AG	TAAGGAGAGAACTACTATTGATTAGAGCTTAACCCAGGTTAACTGCAAGAAGAGG	2097
Db	577	AG	TAAGGAGAGAACTACTATTGATTAGAGCTTAACCCAGGTTAACTGCAAGAAGAGG	518
QY	2098	CG	GAATCTTCCAGCTTTCCTACTGATGCTAATAAGCCAATGTAGTCGAGTTCTTA	2157
Db	517	CG	GAATCTTCCAGCTTTCCTACTGATGCTAATAAGCCAATGTAGTCGAGTTCTTA	458

QY	2158	AGATCATGTTCCAAAGCTAACTGAATCCCACTTCCAATACACACTCATGAACCTCCTGATGGA	2217
Db	457	AGATCATGTTCCAAAGCTAACTGAATCCCACTTCCAATACACACTCATGAACCTCCTGATGGA	398
QY	2218	ACAATAACAGGCCCAAGCCGTGTGTATGATGTGCACACTTGTAGACTCAGAAAAATAC	2277
Db	397	ACAATAACAGGCCCAAGCCGTGTGTATGATGTGCACACTTGTAGACTCAGAAAAATAC	338
QY	2278	TACTCTCATAAATGGGTGGGAGTATTTTGGTGCACCACTACTTTGCTTGGCTGAGTGAAG	2337
Db	337	TACTCTCATAAATGGGTGGGAGTATTTTGGTGCACCACTACTTTGCTTGGCTGAGTGAAG	278
QY	2338	GAATGATATTTCATATATTCATTTATTCATGGACATTTAGTTAGTGTCTTTTATATACCA	2397
Db	277	GAATGATATTTCATATATTCATTTATTCATGGACATTTAGTTAGTGTCTTTTATATACCA	218
QY	2398	GGCATGATGCTGAGTGACACTCTTGTGTATATTTCCAAATTTTTTGTATAGTCGCTGCACA	2457
Db	217	GGCATGATGCTGAGTGACACTCTTGTGTATATTTCCAAATTTTTTGTACAGTCGCTGCACA	158
QY	2458	TATTTGAAATCAAAATATTAAAGCTTTCCAAAATTTGGTCCCTGTTTTTCATGGCAAC	2517
Db	157	TATTTGAAATC-ATATATTAAAGACTTCCWAAGATGAGGTCCCTGGTTTTTCATGGCAAC	99
QY	2518	TTGATCAGTAAGGATTTCCCTCTGTTTGGCACTAAACCATTTACTATATGTTAGACAA	2577
Db	98	TTGATCAGTAAGGATTTCCCTCTGTTTGTACTAAACCATCTACTATGTTTAGACAT	39
QY	2578	GACATTTTTTTTTTTTCCCTCCTCGAAAAAAAATGAGG	2615
Db	38	GACATTTTTTTTTTCTCCTTCYANNDDATAAAAKWDTKD	1

## RESULT 5

BX379978/c	1052 bp	linear	EST 08-MAY-2003
LOCUS	1052 bp	linear	EST 08-MAY-2003
DEFINITION	Hom sapiens p1ACENT COT 25-NORMALIZED Homo sapiens cDNA clone CS001042Y121 3-PRIME, mRNA sequence.		
ACCESSION	BX379978		
VERSION	1		
KEYWORDS	GI:30454806		
SOURCE	EST.		
ORGANISM	Hom sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.		
TITLE	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage Bp 191 91006 EVRY cedex - France Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4396.f For more information about this cluster, see http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS001042AE11NP1&cluster=4396.f. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS001042AE11NP1. Location/Qualifiers 1. 1052		
FEATURES			
source			

FEATURES  
SOURCE

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI042YI21"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-colligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

## ORIGIN

Query Match 38.1%; Score 1014.6; DB 13; Length 1052;  
Best Local Similarity 98.0%; Pred. No. 4.1e-154;  
Matches 1031; Conservative 3; Mismatches 17; Indels 1; Gaps 1;

QY 1539 GCTTGGCCATATTTGTCACCTGATCCCTCTTGGTGTCACAAAAACACAGGAATACA 1598  
Db 1052 GCTTGGCCATATTTGTCACCTGATCCCTCTTGGTGTCACAAAAACACAGGAATACA 993  
QY 1599 ACCCAATAGAAAATAGTCTCGGAAATGTGTGAGAGCAAAAGGCTGAGTCTCTTCTCA 1658  
Db 992 ACCCAATAGAAAATAGTCTCGGAAATGTGTGAGAGCAAAAGGCTGAGTCTCTTCTCA 933  
QY 1659 ACCGTGCAAAAGCGGTGTTCTCCGGGAAACACAGGAAAGATCCGCTACTCAAAACC 1718  
Db 932 ACCGTGCAAAAGCGGTGTTCTCCGGGAAACACAGGAAAGATCCGCTACTCAAAACC 873  
QY 1719 AAGAAATTTAAAGAGGTTTCTTAAATTTTCGACCTTGTCTGAAGCTCACTTTTCAGTGCC 1778  
Db 872 AAGAAATTTAAAGAGGTTTCTTAAATTTTCGACCTTGTCTGAAGCTCACTTTTCAGTGCC 813  
QY 1779 ATTGATGTCAGATGCTGAGTGGCTATTAACTTTTTCCTAAAGATTTAGTTTAAA 1838  
Db 812 ATTGATGTCAGATGCTGAGTGGCTATTAACTTTTTCCTAAAGATTTAGTTTAAA 753  
QY 1839 TAGATATTGTTTGGGAAAGTGAATTTTATAGTTTAAATGCTCATTTTAGAGTGG 1898  
Db 752 TAGATATTGTTTGGGAAAGTGAATTTTATAGTTTAAATGCTCATTTTAGAGTGG 693  
QY 1899 GGAGAGGATATATCTGAGGAGCTTCCAGGCATGTTGGAACCTGATAAAAGCAACTTA 1958  
Db 692 GGAGAGGATATATCTGAGGAGCTTCCAGGCATGTTGGAACCTGATAAAAGCAACTTA 633  
QY 1959 GCAAGGCTCTTTTCATTTATTTTATGTTTTCACCTTAAAGCTTTAGTAACTAGTAGG 2018  
Db 632 GCAAGGCTCTTTTCATTTATTTTATGTTTTCACCTTAAAGCTTTAGTAACTAGTAGG 573  
QY 2019 ATAGAAAACACTGTGCTCCGAGAGTAAAGAGAGCTACTATTGATTAGAGCTTAACCA 2078  
Db 572 ATAGAAAACACTGTGCTCCGAGAGTAAAGAGAGCTACTATTGATTAGAGCTTAACCA 513  
QY 2079 GGTAACTGCAAGAGAGCGGATCTTTCAGCTTTCCATGTAAGCTGATGATGATGATGATG 2138  
Db 512 GGTAACTGCAAGAGAGCGGATCTTTCAGCTTTCCATGTAAGCTGATGATGATGATGATGATG 453  
QY 2139 CAATGTAGTCCAGTTTCTAAGATCATGTTCCAGCTAACTGAATCCCACTTCAATACACA 2198  
Db 452 CAATGTAGTCCAGTTTCTAAGATCATGTTCCAGCTAACTGAATCCCACTTCAATACACA 393  
QY 2199 CTCATGAATCTGATGGAACATATACAGGCCCAAGCTGTGATGATGATGATGATGATGATG 2258  
Db 392 CTCATGAATCTGATGGAACATATACAGGCCCAAGCTGTGATGATGATGATGATGATGATG 333  
QY 2259 CTAGACTCAGAAAAATATCTCTCATAAATGGGTGGAGATTTTGGTGACAACTAC 2318  
Db 332 CTAGACTCAGAAAAATATCTCTCATAAATGGGTGGAGATTTTGGTGACAACTAC 273  
QY 2319 TTTGCTTGGCTGAGTGAAGATGATATTATATATATATATATATATATATATATATATATAT 2378  
Db 272 TTTGCTTGGCTGAGTGAAGATGATATTATATATATATATATATATATATATATATATATAT 213  
QY 2379 TAGTGTCTTTTATATACAGGATGATGCTGAGTGACCTCTGTGTATATATATATATATATAT 2438  
Db 212 TAGTGTCTTTTATATACAGGATGATGCTGAGTGACCTCTGTGTATATATATATATATAT 153  
QY 2439 TTTGTATAGTCGCTGCACATATTTGAAATCAAAATATTAAGACTTTCCAAAAATTTGGTC 2498  
Db 152 TTTGTATAGTCGCTGCACATATTTGAAATC-ATATATTAAGACTTTCCAAAGATGAGTGC 94  
QY 2499 CTTGGTTTTCATGGCAACTGATAGTAAGATTTCCCTCTGTTTGGAACTTAACCA 2558  
Db 93 CTTGGTTTTCATGGCAACTGATAGTAAGATTTCCCTCTGTTTGGAACTTAACCA 34

QY 2559 TTTACTATATGTTAGACAGACATTTTTTTTT 2590  
Db 33 TCTACTTATGTTGACATGACATTCCTTTTT 2

RESULT 6  
AK079220 2431 bp mRNA linear HTC 19-SEP-2003  
LOCUS Mus musculus adult male urinary bladder cDNA, RIKEN full-length  
DEFINITION enriched library, clone:953003820 product:glycoprotein  
(transmembrane) nmb, full insert sequence.  
ACCESSION AK079220  
VERSION AK079220.1 GI:26347858  
KEYWORDS HTC; CAP trapper.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 Carninci, P. and Hayashizaki, Y.  
TITLE High-efficiency full-length cDNA cloning  
JOURNAL Mech. Enzymol. 303, 19-44 (1999)  
MEDLINE 99279253  
PUBMED 10349636  
REFERENCE 2  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)  
JOURNAL 20499374  
MEDLINE 11042159  
PUBMED 11042159  
REFERENCE 3  
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
Konno, H., Akiyama, J., Nishi, K., Kusunai, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,  
Yanai, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, B., Watahiki, M.,  
Yoneda, Y., Iehikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multichannel sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
JOURNAL 20530913  
MEDLINE 11076861  
PUBMED 11076861  
REFERENCE 4  
The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)  
JOURNAL 20530913  
MEDLINE 11076861  
PUBMED 11076861  
REFERENCE 5  
The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
JOURNAL 20530913  
MEDLINE 11076861  
PUBMED 11076861  
REFERENCE 6  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,  
Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N.,  
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akai, S.,  
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome





QY 1421 AACCTACCTGGGGATGACACAGCCTGGCTCTCAGAGCACCCCTGATTTCTGTTCCCT 1480  
DB 1458 AATTTCACCTGGGAGATGATGCAAGCCTGGCCCTCACACACCCCTGATCTATCCCT 1517  
QY 1481 GACAGAGACCCAGCCTGGCCTTTAAGGATGCAACAGAGTGCCTGATCTCCGTTGGCTGC 1540  
DB 1518 GGCAAGAGCCAGACTCCCTCTCAGAGCAGTGAATGGTGTCTGATCTCCATCGGCTGC 1577  
QY 1541 TTGGCCATATTGTCACGTGATCTCCCTCTTGGTGTACAA 1591  
DB 1578 CTGGCTGTGCTGTGTCACCATGTTACCATCTTGTGTACAA 1618

RESULT 7  
BX423077/c  
LOCUS  
DEFINITION BX423077 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone EST 15-MAY-2003  
CS0DN004YA18 3-PRIME, mRNA sequence.

ACCESSION BX423077.1 GI:30766236

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1003)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 4396.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DN004BA09NP1&cluster=4396.f. Contact:

Feng Liang Email: fliang@lifetech.com

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID: CS0DN004BA09NP1.

Location/Qualifiers

1..1003

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DN004YA18"

/tissue\_type="ADULT BRAIN"

/dev\_stage="adult"

/clone\_lib="Homo sapiens ADULT BRAIN"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-strand cDNA was digested with Not I and

cloned into the Not I and EcoRV sites of the pCMVSPORT 6

vector. Library was not normalized."

## ORIGIN

Query Match 35.2%; Score 937; DB 13; Length 1003;  
Best Local Similarity 98.0%; Pred. No. 1.4e-141;  
Matches 968; Conservative 1; Mismatches 17; Indels 2; Gaps 2;  
QY 1598 AACCCAAATAGAAATAGTCTGGGATGTTGGTCAGAGCAAGAGCGCTGAGTGTCTTCTC 1657  
DB 1003 ACCCCAAATAGAAATAGTCTGGGATGTTGGTCAGAGCAAGAGCGCTGAGTGTCTTCTC 945  
QY 1658 AACCGTGCAAAAGCCGTGTTCTTCCCGGGAAACAGGAAAGGATCCGCTACTCAAAAAC 1717  
DB 944 AACCGTGCAAAAGCCGTGTTCTTCCCGGGAAACAGGAAAGGATCCGCTACTCAAAAAC 885  
QY 1718 CAAGAATTTAAAGAGGTTCTTAAATTTCAACCTGTTTCTGAGCTCACTTTTCAGTGC 1777  
DB 884 CAAGAATTTAAAGAGGTTCTTAAATTTCAACCTGTTTCTGAGCTCACTTTTCAGTGC 825

## FEATURES

source

QY 1778 CATTGATGTGAGATGTCTGAGTGGCTATTAACTTTTTTCTCTAAAGATTATTGTAA 1837  
DB 824 CATTGATGTGAGATGTCTGAGTGGCTATTAACTTTTTTCTCTAAAGATTATTGTAA 765  
QY 1838 ATAGATATTGTGTTTGGGGAAGTTGAATTTTTATAGGTTAAATGTCATTTTACAGATG 1897  
DB 764 ATAGATATTGTGTTTGGGGAAGTTGAATTTTTATAGGTTAAATGTCATTTTACAGATG 705  
QY 1998 GGGAGAGGATTATACAGGAGCTTACCCATGTTGTGAAACTGATAAAAGCAACTT 1957  
DB 704 GGGAGAGGATTATACAGGAGCTTACCCATGTTGTGAAACTGATAAAAGCAACTT 645  
QY 1958 AGCAAGGCTTCTTTTCAATTTTTATGTTTCACTTATAAAGTCTTAGGTTAACTAGTAG 2017  
DB 644 AGCAAGGCTTCTTTTCAATTTTTATGTTTCACTTATAAAGTCTTAGGTTAACTAGTAG 585  
QY 2018 GATGAAACACATGTGTCAGAGATGAGGAGAGAGCTACTATTGATTAGAGCTTAAACC 2077  
DB 584 GATGAAACACATGTGTCAGAGATGAGGAGAGAGCTACTATTGATTAGAGCTTAAACC 525  
QY 2078 AGGTTAACTGCAAGAGAGGCGGATACCTTTCAGCTTTCCATGTAACCTGTATGCATAAG 2137  
DB 524 AGGTTAACTGCAAGAGAGGCGGATACCTTTCAGCTTTCCATGTAACCTGTATGCATAAG 465  
QY 2138 CCAATGTAGTCCAGTTCTAAGATCATGTTCCAAAGCTAACTGAATCCCACTTCAATACAC 2197  
DB 464 CCAATGTAGTCCAGTTCTAAGATCATGTTCCAAAGCTAACTGAATCCCACTTCAATACAC 405  
QY 2198 ACTCATGAACTCTGATGGAACAATAACAGGCCCAAGCCTGTGGTATGATGTGCACACTT 2257  
DB 404 ACTCATGAACTCTGATGGAACAATAACAGGCCCAAGCCTGTGGTATGATGTGCACACTT 345  
QY 2258 GCTAGACTCAGAAAAATACTACTCTCATAAATGGTGGGAGTATTTTGGTGACAACCTTA 2317  
DB 344 GCTAGACTCAGAAAAATACTACTCTCATAAATGGTGGGAGTATTTTGGTGACAACCTTA 285  
QY 2318 CTTTCCTGGTGGTGGTGAAGGAATGATATTCATATTAATTTTCCATGTCAGCATTTAG 2377  
DB 284 CTTTCCTGGTGGTGGTGAAGGAATGATATTCATATTAATTTTCCATGTCAGCATTTAG 225  
QY 2378 TTAGTGCCTTTTATATACAGGCATGATGCTGAGTGACACTCTCTGTATATTTTCCAAAT 2437  
DB 224 TTAGTGCCTTTTATATACAGGCATGATGCTGAGTGACACTCTCTGTATATTTTCCAAAT 165  
QY 2438 TTTTGTATAGTCGCTGCACATATTTGAAATCAAAATATTAAGACTTTCCAAATTTTGT 2497  
DB 164 TTTTGTATAGTCGCTGCACATATTTGAAATC-ATATATTAAGACTTTCCAAAGATGAGGT 106  
QY 2498 CCTCGTGTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTTGGAACTAAACCC 2557  
DB 105 CCTCGTGTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTTGGAACTAAACCC 46  
QY 2558 ATTACTATATGTTAGACAGACATTTT 2585  
DB 45 ATCTACTATATGTTATAAANAAGCTATT 18

## RESULT 8

BX458448/c

LOCUS

DEFINITION

3-PRIME, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1015)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

BX458448 1015 bp mRNA linear EST 22-MAY-2003  
BX458448 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE003Y118  
3-PRIME, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1015)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)





Matches	949;	Conservative	6;	Mismatches	18;	Indels	3;	Gaps	2;
Qy	897	ACTACAGTGGAGCTTCGGGATATATCTGCGCTGTTTGTTCACCAATCATCTGTGA	956						
Db	57	AATTCCCGGATGCTTCGGGATATATCTGCGCTGTTTGTTCACCAATCATCTGTGA	116						
Qy	957	ATCACAGTATGTCTCAATGGAACCTTCACGCTTAACCTCACTGTGAAAGCTGCAGAC	1016						
Db	117	ATCACAGTATGTCTCAATGGAACCTTCACGCTTAACCTCACTGTGAAAGCTGCAGAC	176						
Qy	1017	CAGGACCTTTCGCGCACCGCCACCCACCCACCCAGACCTTCAAAACCCACCCCTTTTAG	1076						
Db	177	CAGGACCTTTCGCGCACCGCCACCCACCCACCCAGACCTTCAAAACCCACCCCTTTTAG	236						
Qy	1077	GACCTGCTGGTGACAAACCCCTGAGCTGAGTAGAGTCTCTGATGAAACTGCCAGATTA	1136						
Db	237	GACCTGCTGGTGACAAACCCCTGAGCTGAGTAGAGTCTCTGATGAAACTGCCAGATTA	296						
Qy	1137	ACAGATATGGCAGCTTTCAGGCCACCATCAAAATTTGTAGAGGATCTTAGAGGTTAACA	1196						
Db	297	ACAGATATGGCAGCTTTCAGGCCACCATCAAAATTTGTAGAGGATCTTAGAGGTTAACA	356						
Qy	1197	TCATCCAGATCAGACAGCTCTGATGCGGTGCCATGCCCTGAAAGCTCCCTAATAGACT	1256						
Db	357	TCATCCAGATCAGACAGCTCTGATGCGGTGCCATGCCCTGAAAGCTCCCTAATAGACT	416						
Qy	1257	TTGTGCTGACCTGCCAAGGAGCATTCACGAGGCTCTGACCATCATTTCTGACCCCA	1316						
Db	417	TTGTGCTGACCTGCCAAGGAGCATTCACGAGGCTCTGACCATCATTTCTGACCCCA	476						
Qy	1317	CCTGGAGATCAACCCAGAACACAGTCTGAGCCCTGTGGATGTGGATGAGATGTCTGTC	1376						
Db	477	CCTGGAGATCAACCCAGAACACAGTCTGAGCCCTGTGGATGTGGATGAGATGTCTGTC	536						
Qy	1377	TGACTGTGAGAGCAACCTTCAATGAGTGTGGAGAGTACTGTGTGAACCTCAACCTGGGG	1436						
Db	537	TGACTGTGAGAGCAACCTTCAATGAGTGTGGAGAGTACTGTGTGAACCTCAACCTGGGG	596						
Qy	1437	ATGACACAGCTGCTCTCAGGACACCTGATTTCTGTTCTCCTGACAGACCCAGCCT	1496						
Db	597	ATGACACAGCTGCTCTCAGGACACCTGATTTCTGTTCTCCTGACAGACCCAGCCT	656						
Qy	1497	GCGCTTTAAGGATGCAACACAGTGCCTGATCTCGTTGGCTGTGGCCATATTGTCA	1556						
Db	657	GCGCTTTAAGGATGCAACACAGTGCCTGATCTCGTTGGCTGTGGCCATATTGTCA	716						
Qy	1557	CTGTGATCTCCCTCTTTGGTGTACAAAAACAAAGGAATACAAACCAATAGAAATAGTC	1616						
Db	717	CTGTGATCTCCCTCTTTGGTGTACAAAAACAAAGGAATACAAACCAATAGAAATAGTC	776						
Qy	1617	CTGGGAATGTGTCAAGAGCAAGGCTGAGTGTCTTCTCAACCGTCGAAAGCCGCT	1676						
Db	777	CTGGGAATGTGTCAAGAGCAAGGCTGAGTGTCTTCTCAACCGTCGAAAGCCGCT	836						
Qy	1677	TCTTCCCGGAAACAGGAAAGGATCCGCTACTCAAAACCAAGAAATTAAGAGGTTT	1736						
Db	837	TCTTCCCGGAAA - CAGGAAGGATCCGCTACTCAAAACCAAGAAATTAAGAGGTTT	895						
Qy	1737	CTTAAATTCGACCTTGTTCGAAAGCTCACTTTTCAGTGCCATTTGATGTGAGTGTGCT	1796						
Db	896	CTTAAATTCGACCTTGTTCGAAAGCTCACTTTTCAGTKCCATTTGATGTGAGAGTGCK	955						
Qy	1797	GGAGTGGCTATTAACTTTTCTTAAAGATTATTGTTAAATAGATATTGTTGGTTGG	1856						
Db	956	GGAGTGGCTATTAACTTTTCTTAAAGATTATTGTTAAATAGATATTGTTGGTTGG	1013						
Qy	1857	GAAGTTGAATTTTTTA 1872							
Db	1014	GAGTTGATTTTTTTA 1029							

```
Db 475 TTCCACGAGGCTCTGATCATCTTCTGACCCACCTGCGGATCACCAGACACAG 534
QY 1341 TCTGAGCCCTCTGGATGGATGAGATGTGTCTGCTGCTGAGACGAACTTCAATG 1400
Db 535 TCTGAGCCCTCTGGATGGATGAGATGTGTCTGCTGCTGAGACGAACTTCAATG 594
QY 1401 GGTCTGGGAGCTACTGTGTGAACCTCACCTCGGGGATGACACAGCTGGCTCTCACGA 1460
Db 595 GGTCTGGGAGCTACTGTGTGAACCTCACCTCGGGGATGACACAGCTGGCTCTCACGA 654
QY 1461 GCACCTGATTTCTGTCTCTGCTGACAGAGACCCAGCTCTGCTTTAAGGATGGCAACAGTG 1520
Db 655 GCACCTGATTTCTGTCTCTGCTGACAGAGACCCAGCTCTGCTTTAAGGATGGCAACAGG 714
QY 1521 CCTGATCTCCGTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTA 1580
Db 715 CCTGATCTCCGTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTA 774
QY 1581 AAAACACAAGAAATACAAACCAATAGAAATAGTCTCGGGGAAACAGGAAAGG 1700
Db 775 AAAACACAAGAAATACAAACCAATAGAAATAGTCTCGGGGAAACAGGAAAGG 894
QY 1641 GCCTGAGTGTCTTCTCAACCGTGCAGAAAGCGCTGTCTTCCCGGGAACAGGAAAGG 1700
Db 835 GCCTGAGTGTCTTCTCAACCGTGCAGAAAGCGCTGTCTTCCCGGGAACAGGAAAGG 894
QY 1701 ATCCGCTACTCAAAACCAAGAAATTAAGAGATTTCTTAATTCGACCTTGTTCGA 1760
Db 895 ATCCGCTACTCAAAACCAAGAAATTAAGAGATTTCTTAATTCGACCTTGTTCGA 954
QY 1761 AGCTCACATTTTCAGTGCCATGATGTGAGATGTCTGGAGTGGCTATTAACCTTTTTC 1820
Db 955 AGCTCACATTTTCAGTGCCATGATGTGAGATGTCTGGAGTGGCTATTAACCTTTTTC 1820
QY 1821 CTAAGATTAATGTTAAATAGATATTTGTGTTTGGGGAAG 1860
Db 1015 CTAAGATTAATGTTAAATAGATATTTGTGTTTGGGGAAG 1054
```

## RESULT 11

```
EX336884 1201 bp mRNA linear EST 01-MAY-2003
LOCUS EX336884 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CSODI033YF05 5-PRIME, mRNA sequence.
ACCESSION EX336884
VERSION EX336884.1 GI:30308517
KEYWORDS EST.
SOURCE Homo sapiens (human)
```

## ORGANISM

```
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

## REFERENCE

```
1. (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
```

## AUTHORS

```
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 131 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4396.f For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSODI033CC03Qp1
&cluster=4396.f. Contact : Feng Liang Email : fliang@lifetech.com
URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODI033CC03Qp1.
Location/Qualifiers
```

## FEATURES

source

```
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI033YF05"
```

```
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
```

## ORIGIN

```
Query Match 34.2%; Score 910.6; DB 13; Length 1201;
Best Local Similarity 99.2%; Pred. No. 2.2e-137;
Matches 935; Conservative 1; Mismatches 5; Indels 2; Gaps 2;
QY 28 AACCTTGGTGGCTCGCTCGGTGAGAAATTCAGATGGAATGCTCTACTATTTCTGGGAT 87
Db 103 ACCTTGAGTGGCTCGCTCGGTGAGAAATTCAGATGGAATGCTCTACTATTTCTGGGAT 162
QY 88 TTCTGCTCTGGCTGCAAGATGCGATGCGCCAAACGATTCATGATGTCGTGG 147
Db 163 TTCTGCTCTGGCTGCAAGATGCGATGCGCCAAACGATTCATGATGTCGTGG 222
QY 148 GCAATGAAAGACCTTCTGCTTACATGAGGAGACCAATCAATTAATGCTGCTTCTG 207
Db 223 GCAATGAAAGACCTTCTGCTTACATGAGGAGACCAATCAATTAATGCTGCTTCTG 282
QY 208 ATGAAATGACTGGAATGAAATACTCTACCCAGTGTGGAAGCGGGAGACATGAGTGGA 267
Db 283 ATGAAATGACTGGAATGAAATACTCTACCCAGTGTGGAAGCGGGAGACATGAGTGGA 342
QY 268 AAAAATCTCTGAAAGGAGCGCTGTGAGGCGGTCTGACAGTACTCACCACCTCG 327
Db 343 AAAAATCTCTGAAAGGAGCGCTGTGAGGCGGTCTGACAGTACTCACCACCTCG 402
QY 328 TGGGCTCAAAATATAACATTTTCGCTGAACCTGATTTCCCTAGATGCGAAAGAGATG 387
Db 403 TGGGCTCAAAATATAACATTTTCGCTGAACCTGATTTCCCTAGATGCGAAAGAGATG 462
QY 388 CCAATGCAACATAGTCTATGAGAGAACTGCAGAAATGAGGCTGTTTATCTGCTGATC 447
Db 463 CCAATGCAACATAGTCTATGAGAGAACTGCAGAAATGAGGCTGTTTATCTGCTGATC 522
QY 448 CATATGTTTACACTGACAGCATGTTGTCAGAGGACAGTACGCGGGAATAAGCACCGGCC 507
Db 523 CGTATGTTTACACTGACAGCATGTTGTCAGAGGACAGTACGCGGGAATAAGCACCGGCC 582
QY 508 AAAGCCATCATAACTCTCCCTGATGGGAAACCTTTTCTCACCACCGGATGGAGAA 567
Db 583 AAAGCCATCATAACTCTCCCTGATGGGAAACCTTTTCTCACCACCGGATGGAGAA 642
QY 568 GATGGAATTTTCACTACGCTTCTCCACACACTTGGTCAGTATTTCCAGAAATTTGGAGCAT 627
Db 643 GATGGAATTTTCACTACGCTTCTCCACACACTTGGTCAGTATTTCCAGAAATTTGGAGCAT 702
QY 628 GTTCAGTGAGAGTCTCTGTAACACAGCAAGCAATGTCACCTTGGGCTCAACTCATGGAAG 687
Db 703 GTTCAGTGAGAGTCTCTGTAACACAGCAAGCAATGTCACCTTGGGCTCAACTCATGGAAG 762
QY 688 TGACTGTCTACAGAGACATGAGCGGCATATGTCCTCCATCGCACAGTGAAGATGTGT 747
Db 763 TGACTGTCTACAGAGACATGAGCGGCATATGTCCTCCATCGCACAGTGAAGATGTGT 822
QY 748 ACCTGTAACAGATCAGATTTCTGTTTGTGTAATGTTTCCAGAGAAACGATCGAAAT 807
Db 823 ACCTGTAACAGATCAGATTTCTGTTTGTGTAATGTTTCCAGAGAAACGATCGAAAT 882
QY 808 CATCCGACCAACCTTCC-CAAGATCTCCCATATGTTTGTGATGTCCTGATTCATGATC 866
Db 883 CATCCGACCAACCTTCC-CAAGATCTCCCATATGTTTGTGATGTCCTGATTCATGATC 942
QY 867 CTAGCCACTTCTCAATTAATTTTACCATTAACTACAGTGGAGCTTCGGGGATAATCTG 926
Db 943 CTAGCCACTTCTCAATTAATTTTACCATTAACTACAGTGGAGCTTCGGGGATAATCTG 1002
QY 927 GCCTGTTTGTTCACCAATCATACTATGATGATCATCACGATATGT 969
```

```
|||||
1003 GCCTGTTGTTT-CMCAATCATACTGTGAATCACATGATGT 1044

RESULT 12
BX364871 1201 bp mRNA linear EST 05-MAY-2003
LOCUS BX364871 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CSOD1024YA06 5-PRIME, mRNA sequence.
ACCESSION BX364871
VERSION BX364871.1 GI:30380860
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4396.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS1A1006ZG04QP1&cluster=4396.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS1A1006ZG04QP1.
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1024YA06"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source
Query Match 34.1%; Score 906.4; DB 13; Length 1201;
Best Local Similarity 93.8%; Pred. NO. 1.1e-136;
Matches 959; Conservative 23; Mismatches 34; Indels 6; Gaps 4;

QY 28 AACCTTGGTCCCTGCGTCCGTGAGAAATTCAGCATGGAATGCTCTACTATTTCCTGGAT 87
DB 94 ACCTTGAGTGCCTGCGTCCGTGAGAAATTCAGCATGGAATGCTCTACTATTTCCTGGAT 153
QY 88 TTCTGCTCTGCTGCAAGATTGCCACTTGATCCGCCCAACGATTTCATGATGCTGG 147
DB 154 TTCTGCTCTGCTGCAAGATTGCCACTTGATCCGCCCAACGATTTCATGATGCTGG 213
QY 148 GCAATGAAGACCTTCTGCTTACATGAGGAGCACAAATCAATTAATGGCTGCTCTTG 207
DB 214 GCAATGAAGACCTTCTGCTTACATGAGGAGCACAAATCAATTAATGGCTGCTCTTG 273
QY 208 ATGAAATGACTGAATGAAAACCTCTACCCAGTGGAGCGGGAGACATCAGGTGGA 267
DB 274 ATGAAATGACTGGERATGAAAACCTCTACCCAGTGGAGCGGGAGACATCAGGTGGA 333
QY 268 AAAACTCTCTGAAGGAGGCGGTGTGTCAGCGGTCTCTGACAGTACTACCGCCCTCG 327
DB 334 AAAACTCTCTGAAGGAGGCGGTGTGTCAGCGGTCTCTGACAGTACTACCGCCCTCG 393
QY 328 TGGGCTCAATATATACATTTGGCGGTGAACCTGATATTCCTAGATGCAAAAGGAATG 387
DB 394 TGGGCTCAATATATACATTTGGCGGTGAACCTGATATTCCTAGATGCAAAAGGAATG 453

ORIGIN
Query Match 34.1%; Score 906.4; DB 13; Length 1201;
Best Local Similarity 93.8%; Pred. NO. 1.1e-136;
Matches 959; Conservative 23; Mismatches 34; Indels 6; Gaps 4;

QY 28 AACCTTGGTCCCTGCGTCCGTGAGAAATTCAGCATGGAATGCTCTACTATTTCCTGGAT 87
DB 94 ACCTTGAGTGCCTGCGTCCGTGAGAAATTCAGCATGGAATGCTCTACTATTTCCTGGAT 153
QY 88 TTCTGCTCTGCTGCAAGATTGCCACTTGATCCGCCCAACGATTTCATGATGCTGG 147
DB 154 TTCTGCTCTGCTGCAAGATTGCCACTTGATCCGCCCAACGATTTCATGATGCTGG 213
QY 148 GCAATGAAGACCTTCTGCTTACATGAGGAGCACAAATCAATTAATGGCTGCTCTTG 207
DB 214 GCAATGAAGACCTTCTGCTTACATGAGGAGCACAAATCAATTAATGGCTGCTCTTG 273
QY 208 ATGAAATGACTGAATGAAAACCTCTACCCAGTGGAGCGGGAGACATCAGGTGGA 267
DB 274 ATGAAATGACTGGERATGAAAACCTCTACCCAGTGGAGCGGGAGACATCAGGTGGA 333
QY 268 AAAACTCTCTGAAGGAGGCGGTGTGTCAGCGGTCTCTGACAGTACTACCGCCCTCG 327
DB 334 AAAACTCTCTGAAGGAGGCGGTGTGTCAGCGGTCTCTGACAGTACTACCGCCCTCG 393
QY 328 TGGGCTCAATATATACATTTGGCGGTGAACCTGATATTCCTAGATGCAAAAGGAATG 387
DB 394 TGGGCTCAATATATACATTTGGCGGTGAACCTGATATTCCTAGATGCAAAAGGAATG 453
```

http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0DE003BE09QPl&cluster=4396.f. Contact :  
 Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0DE003BE09QPl.  
 Location/Qualifiers  
 1..1012

## FEATURES

source

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DE003YI18"  
 /tissue\_type="PLACENTA"  
 /clone\_lib="Homo sapiens PLACENTA"  
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
 with a NotI-oligo(dT) primer. Five prime end enriched,  
 double-strand cDNA was digested with Not I and cloned into  
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
 Library was not normalized."

## ORIGIN

Query Match 33.3%; Score 885.6; DB 13; Length 1012;  
 Best Local Similarity 99.2%; Pred. No. 2.6e-133;  
 Matches 899; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

28 AACCTTGGTGGCTCGCTCGGTCGAGATTTCAGCATGGAAATGCTCTACTATTTCCTGGAT 87  
 107 ACCTTGAGTGCTCGCTCGGTCGAGAAATTCAGCATGGAAATGCTCTACTATTTCCTGGAT 166  
 88 TTCTGCTCTGGCTGCGAGAGTTGCCACTGTGATGCGGCAACAGATTTCATGATGCTGG 147  
 167 TTCTGCTCTGGCTGCGAGAGTTGCCACTGTGATGCGGCAACAGATTTCATGATGCTGG 226  
 148 GCATGAAGACCTTCTGCTTACATGAGGAGACATCAATTAATTAATTAATTAATTAAT 207  
 227 GCATGAAGACCTTCTGCTTACATGAGGAGACATCAATTAATTAATTAATTAATTAAT 286  
 208 ATGAAATGATGCTGAATGAATAACTCTACCCAGTGTGAGCGGAGAGATGAGTGA 267  
 287 ATGAAATGATGCTGAATGAATAACTCTACCCAGTGTGAGCGGAGAGATGAGTGA 346  
 268 AAACTCTGAGAGGAGCGGTGTGAGCGGTCCTGACCACTGACTCACCAGCCCTCG 327  
 347 AAACTCTGAGAGGAGCGGTGTGAGCGGTCCTGACCACTGACTCACCAGCCCTCG 406  
 328 TGGCTCAATATATACATTTCCGCTGAACTCTATTTCCCTAGATGCGAAAGAGATG 387  
 407 TGGCTCAATATATACATTTCCGCTGAACTCTATTTCCCTAGATGCGAAAGAGATG 466  
 388 CCAATGGCAACATAGTCTATGAGAGAACTGCAGAAATGAGCTGTTTATCTGCTGATC 447  
 467 CCAATGGCAACATAGTCTATGAGAGAACTGCAGAAATGAGCTGTTTATCTGCTGATC 526  
 448 CATATGTTTACATCTGACAGCATGTGTACAGAGACAGTGCAGGAGAAATGCGCCGCC 507  
 527 CGTATGTTTACATCTGACAGCATGTGTACAGAGACAGTGCAGGAGAAATGCGCCGCC 586  
 508 AAAGCCATCATACGCTTCCCTGATGGGAACTTCTTCTACACCCCGATGGAGAA 567  
 587 AAAGCCATCATACGCTTCCCTGATGGGAACTTCTTCTACACCCCGATGGAGAA 646  
 568 GATGGAATTTATCTACCTGCTTCCACACACTTGGTCACTATGCTATTTCCAGAAATTCGGACGAT 627  
 647 GATGGAATTTATCTACCTGCTTCCACACACTTGGTCACTATTTCCAGAAATTCGGACGAT 706  
 628 GTTCACTGAGAGTTTCTGTGAACACAGCCCAATGTGACACTTGGGCTCAACTCATGGAAG 687  
 707 GTTCACTGAGAGTTTCTGTGAACACAGCCCAATGTGACACTTGGGCTCAACTCATGGAAG 766  
 688 TGACTCTTACAGACATGAGCGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT 747  
 767 TGACTCTTACAGACATGAGCGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT 826  
 748 ACGTGTAAACAGATCAGATTCTCTGTTGTGACTATGTTCCAGAGAACGATCGAAAT 807

Db 827 ACCTGTTAACAGATCAGATTCTGCTGTTGTGACTATGTTCCAGAGAACGATCGAAAT 886  
 QY 808 CATCGAGAAACCTTCC-CAAAGATCTCCCATTTATTTGATGTCCTGATTCATGATC 866  
 Db 887 CATCGAGAAACCTTCTCTCAAGATCTCCCATTTATTTGATGTCCTGATTCATGATC 946  
 QY 867 CTAGCCACTTCTCAATTAATTTCTACCAATTAACCAAGTGGAGCTTCGGGATATACTG 926  
 Db 947 CTAGCCACTTCTCAATTAATTTCTACCAATTAACCAAGTGGAGCTTCGGGATATACTG 1006  
 QY 927 GCCTGT 932  
 Db 1007 GCCTGT 1012

RESULT 14  
BX406948/c

LOCUS BX406948 1138 bp mRNA linear EST 15-MAY-2003  
 DEFINITION BX406948 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone  
 CS0DN005YJ09 3-PRIME, mRNA sequence.  
 ACCESSION BX406948  
 VERSION BX406948.1 GI:30766733  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1138)  
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 CONTACT Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 4396.f For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0AN005CE05NP1&cluster=4396.f. Contact :  
 Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0AN005CE05NP1.  
 FEATURES  
 source  
 1..1138  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DN005YJ09"  
 /tissue\_type="ADULT BRAIN"  
 /dev\_stage="adult"  
 /clone\_lib="Homo sapiens ADULT BRAIN"  
 /note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA  
 was primed with a NotI-oligo(dT) primer. Five prime end  
 enriched, double-strand cDNA was digested with Not I and  
 cloned into the Not I and EcoRV sites of the pCMVSPORT 6  
 vector. Library was not normalized."

## ORIGIN

Query Match 32.6%; Score 868; DB 13; Length 1138;  
 Best Local Similarity 94.8%; Pred. No. 1.7e-130;  
 Matches 912; Conservative 13; Mismatches 33; Indels 4; Gaps 3;

QY 1650 TCTTTCTCAACCGTGCAGAAAGCGTGTCTTCCCGGAAACGAGGAAATCCGCTAC 1709  
 Db 962 TTTTTCACCGTGCAGAAAGCGTGT--CTTCCGAAACCGAGGAAATCCGCTAC 905  
 1710 TCAAAAACCAAGATTTTAAAGGAGTTCTTAATTTCCGACCTTCTTGAAGCTCACTT 1769  
 Db 904 TCAAAAACCAAGATTTAAAGGAGTTCTTAATTTCCGACCTTCTTGAAGCTCACTT 845  
 QY 1770 TTCAGTGCAATGATGTGAGATGTGCTGAGTGCTATTAACTTTTCTCTAAAGATT 1829

Db 844 TTKAGTGCCATGATGTGAGATGCTGGAGTGGCTATTAACCTTTTTCCTAAAGATT 785  
QY 1830 ATTGTTAAATAGATATTGTGTTTGGGGAAGTTGAATTTTTTATAGTTAAATGTCAATT 1889  
Db 784 ATTGTTAAATAGATATTGTGTTTGGGGAAGTTGAATTTTTTATAGTTAAATGTCAATT 725  
QY 1890 TAGAGATGGGAGAGGATATATCTGAGGAGCTTTCAGCCATGTTGTGAACCTGATAAA 1949  
Db 724 TAGAGATGGGAGAGGATATATCTGAGGAGCTTTCAGCCATGTTGTGAACCTGATAAA 665  
QY 1950 AGCAACTTAGCAAGCTTCTTTTCATTA-TTTTTATGTTTCACTTATAAGTCTTAGGT 2008  
Db 664 AGCAACTTAGCAAGCTTCTTTTCATTA-TTTTTATGTTTCACTTATAAGTCTTAGGT 605  
QY 2009 AACTAGTAGGATAGAAACACTGTGTCGAGAGTAGGAGAGAGAGCTACTATTGATTAGA 2068  
Db 604 CACTMGTAGGATAGAAACACTGTGTCGAGAGTAGGAGAGAGAGCTACTATTGATTAGA 545  
QY 2069 GCCTAACCCAGGTTAACTGCAAGAGAGGCGGGATCTTTCAGCTTTCATGTAACCTGTA 2128  
Db 544 GCCTAACCCAGGTTAACTGCAAGAGAGGCGGGATCTTTCAGCTTTCATGTAACCTGTA 485  
QY 2129 TGCATAAGCCAAATGTAGTCCAGTTTCTAAGATCATGTTTCAAGCTTAAGTGAATCCCACT 2188  
Db 484 TGCATAAGCCAAATGTAGTCCAGTTTCTAAGATCATGTTTCAAGCTTAAGTGAATCCCACT 425  
QY 2189 TCAATACACACTCATGAATCTCTGTATGAGAAACAATAAGAGGCGGAGGCTGTGTTATGATG 2248  
Db 424 TCAANACANACTCATGGACTCTGTATGGAACAATCACAGGCGGAGGCTGTGTTATGATG 365  
QY 2249 TGCACACTTGTAGACTCAGAAAATAATCTACTCTCATAAATGGGTGGGAGTATTTGGT 2308  
Db 364 TGCACACTTGTAGACTCAGAAAATAATCTACTCTCATAAATGGGTGGGAGTATTTGGT 305  
QY 2309 GACAACTTACTTGTGCTGAGTGAAGGAATGATATTCATATATTCATTATTTCCATG 2368  
Db 304 GACAACTTACTTGTGCTGAGTGAAGGAATGATATTCATATATTCATTATTTCCATG 245  
QY 2369 GACATTTAGTTAGTCTTTTATATACAGGAGATGCTGAGTGAACCTTCTGTGATA 2428  
Db 244 GACATTTAGTTAGTCTTTTATATACAGGAGATGCTGAGTGAACCTTCTGTGATA 185  
QY 2429 TTTCCAAATTTTGTATGCTGCTGCACATATTCGAATCAAAATTAAGACTTTCCAA 2488  
Db 184 TTTCCAAATTTTGTATGCTGCTGCACATATTCGAATCAAAATTAAGACTTTCCAA 126  
QY 2489 AAATTTGTCCTGTTTTCATGGCACTTGATCAGTAAGGATTTCCCTCTGTTTGA 2548  
Db 125 AGATGAGTCCCTGTTTTCATGGCACTTGATCAGTAAGGATTTCCCTCTGTTTGA 66  
QY 2549 ACTAAACCACTTACTATATGTTAGTGAACAAGATTTTTTTTTTTTCTCTCTGTAATAA 2608  
Db 65 ACTAAACCACTTACTATATGTTAGTGAACAAGATTTTTTTTTTTTCTCTCTGTAATAA 6  
QY 2609 AA 2610  
Db 5 NM 4

RESULT 15  
CF552020  
LOCUS  
DEFINITION  
IMAGE:30529955 5', mRNA sequence.  
CF552020  
VERSION  
CF552020.1 GI:34888854  
EST.  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE

CF552020 895 bp mRNA linear EST 22-SEP-2003  
AGENCY: NIH\_MGC\_183 Homo sapiens cdna clone  
IMAGE:30529955 5', mRNA sequence.  
CF552020  
VERSION  
CF552020.1 GI:34888854  
EST.  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE

Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 895)

AUTHORS  
TITLE  
JOURNAL  
COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics / NIH  
National Cancer Institute  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: ggapbs-r@mail.nih.gov  
Tissue Procurement: Dr. Michael Brownstein  
cdna Library Preparation: Invitrogen Corp  
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: NDAM618 row: f column: 12  
High quality sequence stop: 719.  
Location/Qualifiers  
1..895  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30529955"  
/lab\_host="DH10B-Tona (T1 and T5 phage resistant)"  
/note="Organ: Pooled muscle (cardiac and skeletal);  
Vector: pCMV-SPORT6.1; Site 1: EcoRV (destroyed); Site 2:  
NotI; Library is oligo-dT primed and directionally cloned  
(EcoRV site is destroyed upon cloning). Average insert  
size 1.7. Library was constructed by Invitrogen."

FEATURES  
source

ORIGIN

Query Match 32.6%; Score 867.4; DB 14; Length 895;  
Best Local Similarity 99.7%; Pred. No. 2.4e-130;  
Matches 890; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 771 GTGTTTGTGATGTTCCAGAGAACGATCGAATTCATCGAGAACCTTCCCAA 829  
Db 1 GTGTTTGTGATGTTCCAGAGAACGATCGAATTCATCGAGAACCTTCCCAA 60  
QY 830 GATCTCCCCATTTGTTGATGCTCTGATTCATGATCTTAGCCTTCTCAATTTCT 889  
Db 61 GATCTCCCCATTTGTTGATGCTCTGATTCATGATCTTAGCCTTCTCAATTTCT 120  
QY 890 ACCATTAACTCAAGTGGAGCTTCGGGGATATACTGGCCTGTTTGTTCACCAATCAT 949  
Db 121 ACCATTAACTCAAGTGGAGCTTCGGGGATATACTGGCCTGTTTGTTCACCAATCAT 180  
QY 950 ACTGTGAATCACAGTATGTTGCTCAATGGAACTTCAGCCTTAACCTCACTGTGAAGCT 1009  
Db 181 ACTGTGAATCACAGTATGTTGCTCAATGGAACTTCAGCCTTAACCTCACTGTGAAGCT 240  
QY 1010 GCAGCACCAGGACCTTGTCCGCCACCGCCACACACACACACACCTTCAAAACCCACCT 1069  
Db 241 GCAGCACCAGGACCTTGTCCGCCACCGCCACACACACACACCTTCAAAACCCACCT 300  
QY 1070 TCTTTAGGACCTGCTGTTGACAAACCCCTTGGAGCTGAGTAGGATTCCTGATGAATCTGC 1129  
Db 301 TCTTTAGGACCTGCTGTTGACAAACCCCTTGGAGCTGAGTAGGATTCCTGATGAATCTGC 360  
QY 1130 CAGATTAAAGATATGGCCACTTTTCAAGCCACCATCACAAATTTAGAGGGATCTTAGAG 1189  
Db 361 CAGATTAAAGATATGGCCACTTTTCAAGCCACCATCACAAATTTAGAGGGATCTTAGAG 420  
QY 1190 GTTAACATCATCCAGATGACAGAGCTCTGATGCCGGTCCATGGCCTGAAAGCTCCCTA 1249  
Db 421 GTTAACATCATCCAGATGACAGAGCTCTGATGCCGGTCCATGGCCTGAAAGCTCCCTA 480  
QY 1250 ATAGACTTGTGCTGACCTGCCAAGGGAGATTTCCACGAGGCTGTGTACCATCTTCT 1309  
Db 481 ATAGACTTGTGCTGACCTGCCAAGGGAGATTTCCACGAGGCTGTGTACCATCTTCT 540  
QY 1310 GACCCCACTGGAGATCAACCCAGAACACAGTCTGCGAGCCTGTGGATGTGGATG 1369

[illegible]

Search completed: September 10, 2004, 09:57:44  
Job time : 10044 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: September 10, 2004, 13:27:49 ; Search time 218 Seconds  
(without alignments)  
7702.697 Million cell updates/sec

Title: US-10-039-272-1  
Perfect score: 4850  
Sequence: 1 cggcagggggccagagga.....aaaaaaaaaaaaaaaaaaaaa 2661

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 2034082

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+n2p model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool/p/US10039272/runat\_07092004\_144223\_21629/app\_query.fasta\_1.2823  
-DB=SPTREMBL\_25 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -IOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10039272.qcgn\_1\_1.296@runat\_07092004\_144223\_21629 -NCPU=6 -ICPU=3  
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SPTREMBL\_25.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_nhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2983	61.5	572	4 Q8N1A1	Q8n1a1 homo sapien

ID	Q8N1A1	PRELIMINARY;	PRT;	572 AA.
2	2170	44.7	574	11 Q99P91
3	2170	44.7	574	11 Q9QXA0
4	2170	44.7	574	11 Q8BVV9
5	2166	44.7	574	11 Q8BXL4
6	2136	44.0	572	11 Q9QZF6
7	2015	41.5	526	11 Q8BVA0
8	1029	21.2	206	4 Q96F58
9	1022	21.1	206	4 Q8IXJ5
10	644.5	13.3	626	11 Q9CZB2
11	372	7.7	461	6 Q97884
12	271.5	5.6	270	13 Q93391
13	247.5	5.1	236	11 Q9QY67
14	198.5	4.1	423	4 Q8N0W9
15	198.5	4.1	435	4 Q8N3G9
16	192	4.0	397	4 Q8N3R2
17	167	3.4	141	11 Q9QY70
18	151	3.1	354	4 Q8IY46
19	141.5	2.9	906	17 Q8TPY9
20	135	2.8	446	10 Q22458
21	134.5	2.8	1817	17 Q8TI59
22	131	2.7	879	17 Q8PWJ6
23	128.5	2.6	688	17 Q8TR88
24	127.5	2.6	3988	17 Q8TPZ1
25	126	2.6	603	16 Q7U5X8
26	121.5	2.5	525	11 Q99PS8
27	121.5	2.5	1188	17 Q8TPZ7
28	121	2.5	1081	6 Q9BE73
29	120	2.5	766	4 Q82987
30	119.5	2.5	507	16 Q8FV77
31	119.5	2.5	510	11 Q9ESB2
32	119.5	2.5	2016	17 Q8TUS9
33	119	2.5	1948	17 Q8TI51
34	118.5	2.4	2029	17 Q8TI44
35	118	2.4	2869	6 Q28398
36	117	2.4	883	17 Q8TTT9
37	116.5	2.4	703	11 Q8BGU1
38	116	2.4	1698	17 Q8TPK7
39	114.5	2.4	515	11 Q9TPS7
40	114.5	2.4	594	5 Q9VEP4
41	114	2.4	2395	5 Q27167
42	113.5	2.3	525	11 Q99PS6
43	113.5	2.3	525	11 Q9ESB3
44	113.5	2.3	525	11 Q99PS5
45	113	2.3	475	10 Q949Z1

## ALIGNMENTS

## RESULT 1

ID	Q8N1A1	PRELIMINARY;	PRT;	572 AA.
AC	Q8N1A1			
DT	01-OCT-2002 (TrEMBLrel. 22, Created)			
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)			
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)			
DE	Similar to glycoprotein (Transmembrane) nmb.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain, and Lung;			
RA	Strausberg R.;			
RL	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: BC032783; AAH32783.1; -			
DR	GO: GO:0016021; C:integral to membrane; IEA.			
DR	InterPro: IPR000601; PKD.			
DR	Pfam: PF00801; PKD; 1.			
DR	SMART: SM00089; PKD; 1.			
DR	PROSITE: PS50093; PKD; 1.			
KW	Transmembrane.			
SQ	SEQUENCE 572 AA; 63922 MW; 2465C12C2F0F3996 CRC64;			

Alignment Scores:

[illegible]

US-10-039-272-1 (1-2661) x Q8N1A1 (1-572)

QY	60	ATGGAATGCTCTACTATTCTCTGGGATTTCTGCTCTGGCTGCGAAGATTCCTCACTTCAT	119
Db	1	MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuLeuAlaAargLeuProLeuAsp	20
QY	120	GCGCCAAACGATTTTCATGATGCTGGCGCAATGAAGACCTTCCTTACATGAGGAG	179
Db	21	AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu	40
QY	180	CACAATCAATTAATGGCTGGTCTCTTCGATGAAATGACTGGAATGAAATACTTACC	239
Db	41	HisAsnGlnLeuAsnGlyTyrSerSerAspGluAsnAspTyrAsnGluLysLeuTyrPro	60
QY	240	GTGTGGAAGCGGGAGACATAGAGGTGAAATACTCTCGAAGGAGGCCGTGTGAGCGG	299
Db	61	ValTyrLysArgGlyAspMetArgTyrLysAsnSerTyrLysGlyGlyArgValGlnAla	80
QY	300	GTCCTGACCACTGACTCACAGCCCTCGGGCTCAAATATACATTTCGCGTGAACCTG	359
Db	81	ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu	100
QY	360	ATATTCTCTAGATGCCAAAAAGGAGATGCCAAATGGCAACATAGTCTTATGAGAAGACTGC	419
Db	101	IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys	120
QY	420	AGAAATGAGGCTGTTTATCTGCTGATCCATATGTTTACAACTGGACACGANTGGTCAGAG	479
Db	121	ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTyrThrAlaTyrPheGlu	140
QY	480	GACAGTCACGGGGAAATGGCACCGGCCAAAGCCATCATACGCTCTTCCTCATGGGAAA	539
Db	141	AspSerAspGlyGluAsnGlyThrGlyGlnSerHisHisAsnValPheProaspGlyLys	160
QY	540	CCTTTTCTCTACCCCGGATGGAGAAGATGGAATTTTCATCTACGTCTTCCACACATT	599
Db	161	ProPheProHisHisProGlyTyrArgArgTyrAsnPheIleTyrValPheHisThrLeu	180
QY	600	GGTCAGTATTTCCAGAAATTTGGACGATGTTCTAGTGAGAGTTTCTGTGAACACAGCCAAT	659
Db	181	GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn	200
QY	660	GTGACACTGGCGCTCAACTCATGGAAGTGACTCTTACAGAGACATGACGGGCATAT	719
Db	201	ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr	220
QY	720	GTTCCCATCGCACAGTGAAGATGTGTAAGTGTGAACATCAGATTCCTGTGTGTGTG	779
Db	221	ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal	240
QY	780	ACTATGTTCCAGAACGATCGAAATTCATCCGACGAAACCTTCCC- AAAGATCTCCCC	838
Db	241	ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro	260
QY	839	ATTATGTTTGATGTCCTGATTCATGATCCTAGCCACCTCTCTCAATATTTCTTACCATTAA	898
Db	261	IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn	280
QY	899	TACAGTGGAGCTTCGGGGATAACTGGCTGTTGTTTCCACCAATCATACTGTCAAT	958
Db	281	TyrLysTyrPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn	300
QY	959	CACACCTATGTGCTCAATGGAAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCACA	1018

[illegible]

## RESULT 2

ID	Q99P91		PRELIMINARY;	PRT; 574 AA.
AC	Q99P91;			
DT	01-JUN-2001	(TEMBRel. 17, Created)		
DT	01-JUN-2001	(TEMBRel. 17, Last sequence update)		
DT	01-OCT-2003	(TEMBRel. 25, Last annotation update)		
DE	Dendritic cell-associated transmembrane protein (Glycoprotein			
DE	(transmembrane) nmb).			
GN	GPNMB OR DCHIL.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BALB/c;			
EA	Shikano S., Bonkobara M., Zukas P.K., Ariizumi K. ;			
FT	"Molecular Cloning of a Dendritic Cell-Associated Transmembrane			

RT Protein, DC-HIL, that Promotes RGD-Dependent Adhesion of Endothelial  
RT Cells Through Recognition of Heparan Sulfate Proteoglycans."; [PMID: 15500000](#)  
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

```

RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF322054; AAK14240.1; -.
DR EMBL; BC026375; AAB26375.1; -.
DR MGD; MGI:1334765; Gpmmb.
DR GO; GO:0003687; C:integral to plasma membrane; IDA.
DR GO; GO:0005194; F:cell adhesion molecule activity; IDA.
DR GO; GO:0009201; F:heparin binding; IDA.
DR GO; GO:0005178; F:integrin binding; IDA.
DR GO; GO:0007155; F:cell adhesion; IDA.
DR InterPro; IPR000583; GATase_2.
DR InterPro; IPR000601; PKD.
DR Pfam; PF00801; PKD; 1.
DR SMART; SM00089; PKD; 1.
DR PROSITE; PS00443; GATASE_TYPE_II; 1.
DR PROSITE; PS00093; PKD; 1.
KW Transmembrane.
SQ SEQUENCE 574 AA; 63674 MW; B6AE9AC27AB6ACD0 CRC64;

```

Alignment Scores:	
Pred. No.:	3,05e-202
Score:	2170.00
Percent Similarity:	82.63%
Best Local Similarity:	70.53%
Query Match:	44.74%
DB:	11
	11
	2
Length:	574
Matches:	402
Conservative:	69
Mismatches:	83
Indels:	17
Gaps:	2

US-10-039-272-1 (1-2661) x Q99P91 (1-574)

Qy	60	ATGGAAATGCTCTACTATTTCTGGGAATTCGCTCCTCGCTGCAAGATTGCCACTTGAT	119
Db	1	MetGluSerLeuCysGlyValLeuGlyPheLeuLeuAlaAlaGlyLeuProLeuGln	20
Qy	120	GCCGCCAAACGATTTCATGATGCTGTGGCAATGAAGAAGACCTTCTGCTTACATGAGGAG	179
Db	21	AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAsnHisMetArgGlu	40
Qy	180	CACAATCAATTAATGCGCTGCTCTCATGATGAATAACTGCTGAATGAATAAACTTACCCA	239
Db	41	HisAsnGlnLeuArgGlyTyrPheSerSerAspGluAsnGluTyrAspGluHisLeuTyrPro	60
Qy	240	GTGTGGAACGGGGAGACATCAGGTGTGAAAACTCCTGGAAGGGAGCGCTGTGAGGGC	299
Db	61	ValTyrArgArgGlyAspGlyArgTyrPheLysAspSerTyrGluGlyGlyArgValGlnAla	80
Qy	300	GTCTGACCACTGACTACCAAGCCCTCGTGGCTCAAAATATAACATTTTCGGTGAACCTG	359
Db	81	ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu	100
Qy	360	ATATTCCCTAGATGCCAAAGGAAGATCCCAATGGCAACATAGTCTATGAGAGAACTGC	419
Db	101	ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys	120
Qy	420	AGAAATCAGGCTGGTTTATCTGCTGATCCATATGTTTACAACTGGACACGATGTCAGAG	479
Db	121	ArgAsnAspLeuGlyLeuThrSerAspLeuHisValTyrAsnTrpThrAlaGlyAlaAsp	140
Qy	480	GACAGTGACGGGGAAATGGCACCGGCCAAAGCCCATCAACGCTCTCCCTGATGGGAA	539
Db	141	AspGlyAspTrpGluAspGlyThrSerArgSerGlnHisLeuArgPheProAspArgArg	160
Qy	540	CCATTTCTCACACCCCGGATGAGAAAGATGGAATTTTCATCTACGTCTCCACACACTT	599
Db	161	ProPheProArgProHisGlyTyrPheLysLysTrpSerPheValTyrValPheHisThrLeu	180
Qy	600	GGTCAGPATTCAGAATTGGGACGATGTTTCAGTGAGAGTTTCTGTGAAACACAGCCCAAT	659

Db	181	GlyGlnTyrPheGlnLysLeuGlnValArgCysSerAlaArgValSerIleAsnThrValAsn	200
QY	660	GTCACTTGGGCTCAACTCTGAAGTACTGTCTACAGAAGATCGACGGCAAT	719
Db	201	LeuThrAlaGlyProGlnValMetGluValThrValPheArgTyrGlyArgAlaTyr	220
QY	720	GTTCCCATCGCAAGTGAAGATGTGACGTGGTAAACAGATCACATTCTGTGTTTCGTG	779
Db	221	IleProIleSerLysValIysAspValTyrValIleThrAspGlnIleProValPheVal	240
QY	780	ACTATGTTCCAGAAGACGATCGAAATTCATCCGACGAACCTTCCC-AAAGATCTCCCC	838
Db	241	ThrMetSerGlnLysAsnAspArgAsnLeuSerAspGluIlePheLeuArgAspLeuPro	260
QY	839	ATTATGTTTGATGTCCTGATTCATGATCCTACGCCACTTCCTCAATTATTCTACCATTAAC	898
Db	261	IleValPheAspValLeuIleHisAspProSerHisPheLeuAsnAspSerAlaIleSer	280
QY	899	TCAAGTGGAGCTTCGGGGTAATACTGGCTGTTGTTGTTTCCCAATCATACTGTGAAT	950
Db	281	TyrLysTyrAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn	300
QY	959	CACAGTATGTGCTCAATGGACCTTCAGCCTTAACTCACTGCTGAAAGCTGCAGACCA	1018
Db	301	HisThrTyrValLeuAsnGlyThrPheAsnLeuAsnLeuThrValGlnThrAlaValPro	320
QY	1019	GGACCTCTGTCG-----CCACCCGCCACCA	1042
Db	321	GlyProCysProProProSerProSerThrProProProProSerThrProProSerPro	340
QY	1043	CCACCCAGACTT-----TCAAAACCCACCCCTCTTTTAGACCTGCTGGTGAC	1090
Db	341	ProProSerProLeuProThrLeuSerThrProSerProSerLeuMetProThrGlyTyr	360
QY	1091	AACCCCTGGAGCTGCTAGGATTCCTGATGAATACTGCCAGATTACAGATATGCCAC	1150
Db	361	LysSerMetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyr	380
QY	1151	TTTCAAGCCACCATCACAAATTGTAGAGGGAATCTTAGAGGTAAACATCATCCAGATGACA	1210
Db	381	PheArgAlaThrIleThrIleValGluGlyIleLeuGluValSerIleMetGlnIleAla	400
QY	1211	GACGTCTGTAGCCGGTGCCATGCCCTGAAAGCTCCCTAATAGACTTTGCTGTCGCTGC	1270
Db	401	AspValProMetProThrProGlnProAlaAsnSerLeuMetAspPheThrValThrCys	420
QY	1271	CAACGGAGCATTCACCGAGGTCTGTACCATTCATTTGACCCCACTGCGAGATCAACC	1330
Db	421	LysGlyAlaThrProMetGluAlaCysThrIleIleSerAspProThrCysGlnIleAla	440
QY	1331	CAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTCTGCTGACTGTGACAGCA	1390
Db	441	GlnAsnArgValCysSerProValAlaValAspGlyLeuCysLeuLeuSerValArgArg	460
QY	1391	ACCTTCAATGGTCTGGACGTACTGTGTGAACCTCACCTCGGGGATGACACAAGCCCTG	1450
Db	461	AlaPheAsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAspAlaSerLeu	480
QY	1451	GCTCTCACGAGCACCTGATTCTGTTCTCTGCAGAGACCCAGCTCGCTTAAAGGATG	1510
Db	481	AlaLeuThrSerThrLeuIleSerIleProGlyLysAspProAspSerProLeuArgAla	500
QY	1511	GCAACAGTGCCTGATCTCCGTGTGGCTGTTGGCCATATTGTCTCACTGTGATCTCCCTC	1570
Db	501	ValAsnGlyValLeuIleSerIleGlyCysLeuAlaValLeuValThrMetValThrIle	520
QY	1571	TTGGTGTACAAAAACACAAAGGAATACACCAATAGAAAAATAGTCTCGGGAATGTGTC	1630
Db	521	LeuLeuTyrIlysHisLysAlaTyrLysProIleGlyAsnCysProArgAsnThrVal	540
QY	1631	AGAAGCAAAAGGCTGAGTGTCTTCTCAACCGTGCAAAAGCCGTGTTCTTCCCGGAAAC	1690
Db	541	LysGlyIysGlyLeuSerValLeuLeuSerHisLysAlaLysAlaProPhePheArgLysAsp	560

```

QY 1691 CAGGAAAGAGTCCGCTACTCAAAACCAA 1720
Db 561 GlnGluLysAspProLeuGlnAspLys 570

RESULT 3
Q9QXA0 PRELIMINARY; PRT; 574 AA.
AC Q9QXA0;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Putative transmembrane glycoprotein.
GN GPNMB OR NMB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Bachner D., Schroder D., Gross G.;
RT "Isolation and developmental expression analysis of murine nmb.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL, AL251685; CAB65272.1; -.
DR MGD; MGI:1934765; Gpmmb.
DR GO; GO:0005887; C:integral to plasma membrane; IDA.
DR GO; GO:0005194; F:cell adhesion molecule activity; IDA.
DR GO; GO:0008201; F:heparin binding; IDA.
DR GO; GO:0005178; F:integrin binding; IDA.
DR GO; GO:0007155; P:cell adhesion; IDA.
DR InterPro; IPR000583; GATase_2.
DR InterPro; IPR000601; PKD.
DR Pfam; PF00801; PKD; 1.
DR SMART; SM00089; PKD; 1.
DR PROSITE; PS00443; GATASE_TYPE_II; 1.
DR PROSITE; PS00093; PKD; 1.
KW Transmembrane.
SQ SEQUENCE 574 AA; 63681 MW; E0759D7626F0829A CRC64;

Alignment Scores:
Pred. No.: 3,05e-202 Length: 574
Score: 2170.00 Matches: 402
Percent Similarity: 82.63% Conservative: 69
Best Local Similarity: 70.53% Mismatches: 83
Query Match: 44.74% Indels: 17
DB: 11 Gaps: 2

US-10-039-272-1 (1-2661) x Q9QXA0 (1-574)
QY 60 ATGGATGCTCTACTATTCTCGGGATTTCTGCTCGTGGCTGCAAGATTGCCACTTGAT 119
Db 1 MetGluSerLeuCysGlyValLeuGlyPheLeuLeuLeuAlaAlaGlyLeuProLeuGln 20
QY 120 GCCGCCAAAGATTTCATGATGCTGGGCAATCAAGAACCTCTTCTGCTTACATGAGGAG 179
Db 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40
QY 180 CACAATCAATTAATGCTGCTCTCTGATGAAATGCAATGCAATGAAATCACTACCCA 239
Db 41 HisAsnGlnLeuArgGlyTyrSerSerAspGluAsnGluTyrAspGluHisLeuTyrPro 60
QY 240 GTGTGGAAGGGGAGACATGAGTGGAAAACCTCTGGAAGGGAGCGGTGTGCAGCG 299
Db 61 ValTyrArgArgGlyAspGlyArgTyrPheAspSerTyrPheValValGlnAla 80
QY 300 GTCCTGACAGTACTACACAGCCCTCTGGGGCTCAAAATATAACATTTCGGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
QY 360 ATATTCCTAGATGCCAAAAGAGATGCCAATGGCAACATAGTCTATGAGAAAGACTGC 419
Db 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120

420 AGAAATGAGGCTGTTTATCTGTGATCCATATGTTTACAACTGACAGCATGTGTCAG 479
121 ArgAsnAspLeuGlyLeuThrSerAspLeuHisValTyrAsnTyrThrAlaGlyAlaAsp 140
480 GACAGTGACGGGAAATGGACCGCCCAAGCCATCATACGCTCTCCCTGATGGGAAA 539
141 AspGlyAspTyrPheGluAspGlyThrSerArgSerGlnHisLeuArgPheProAspArg 160
540 CCTTTCTCACCACCCCGGATGGAGAGATGGAATTTTCATCTAGCTTCCACACACTT 599
161 ProPheProArgProHisGlyTyrPheValTyrValPheHisThrLeu 180
600 GGTCAAGTATTTCCAGAAATCGGACGATGTCAGTGAGAGTTTCTGTCACACACCCAT 659
181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerAlaArgValSerIleAsnThrValAsn 200
660 GTGACACTTGGCCCTCAACTCATGGAAGTGTCTACAGAGACATGAGCGGCATAT 719
201 LeuThrAlaGlyProGlnValMetGluValThrValPheArgArgTyrGlyArgAlaTyr 220
720 GTTCCATCGCACAGTGAAGATGTGACGTGTGTAACAGATCAGATTCCTGTGTTGTG 779
221 IleProIleSerLysValAspValTyrValIleThrAspGlnIleProValPheVal 240
780 ACTATGTTCCAGAAACGATCGAAATTCATCCGCGAAACCTTCCC-AAAGATCTCCCC 838
241 ThrMetSerGlnLysAsnAspArgAsnLeuSerAspGluIlePheLeuArgAspLeuPro 260
839 ATTATGTTGATGCTCTGATTCATGCTACCTACCTGCTCAATTTATTTCTACCATTA 898
261 IleValPheAspValLeuIleHisAspProSerHisPheLeuAsnAspSerAlaIleSer 280
899 TACAAGTGGAGCTTCGGGATTAATCTGCGCTGTTTGTTCACCAATCATCTGTGAAT 958
281 TyrLysTyrAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300
959 CACAGTATGCTCAATGGAACCTTCAGCTTACCTCACTGCTGTAAGAACTCGACAGCCA 1018
301 HisThrTyrValLeuAsnGlyThrPheAsnLeuAsnLeuThrValGlnThrAlaValPro 320
1019 GGACCTTGTCCGCCACCGCCACCA-----CCACCCAGACCT----- 1054
321 GlyProCysProProProSerProSerThrProProSerProSerThrProProLeuPro 340
1055 -----TCAAAACCCACCCCTTCTTTTGGAGACTGCTGTGAC 1090
341 SerProSerProLeuProThrLeuSerThrProSerProSerLeuMetProThrGlyTyr 360
1091 AACCCCTCGAGTGTAGTAGTTCCTGATGAAACTGCCAGATTACAGATATGGCCAC 1150
361 LysSerMetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyr 380
1151 TTTCAGACCCACCATCAATTTAGAGGGAATCTTAGAGCTTAACATCATCCAGATGACA 1210
381 PheArgAlaThrIleThrIleValGluGlyLeuValSerIleMetGlnIleAla 400
1211 GAGCTCTGATGCGGTGCCATGCGCTGAAAGCTCCCTAATAGACTTTGTGTCACCTGC 1270
401 AspValProMetProThrProGlnProAlaAsnSerLeuMetAspPheThrValThrCys 420
1271 CAAGGAGACATCCCGAGGCTGTACCATCATTTCTACCCACCTCGGAGATCACC 1330
421 LysGlyAlaThrProMetGluAlaCysThrIleSerAspProThrCysGlnIleAla 440
1331 CAGACACACTCTGCAGCCCTGTGATGTGATGTGATGTCTGCTGACTGTGAGACGA 1390
441 GlnAsnArgValCysSerProValAlaValAspGlyLeuCysLeuLeuSerValArgArg 460
1391 ACCTTCAATGGTCTGGGAGCTACTGTGTGAACCTCACCTCGGGGATGACACAGCCTG 1450
461 AlaPheAsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAspAlaSerLeu 480
1451 GCTCTCAGGACCCCTGATTTCTGTTCTCCTCAGACAGACCCAGCCTCGCTTTAAGGATG 1510

```

D <sub>b</sub>	AlaLeuThrSerThrLeuIleSerIleProGlyLysAspProAspSerProLeuArgAla	500
Q <sub>Y</sub>	GCAACAGTGCCCTGCATCTCGTTGGTCTGTGGCCATAATTTCACCTGATCCTCCCTC	1570
D <sub>b</sub>	ValAsnGlyValLeuIleSerIleGlyCysLeuAlaValLeuValThrMetValThrIle	520
Q <sub>Y</sub>	TTCGTGTCAAAAAACAACGAATACAAACCACATAGAAATAGTCTCGGATGTGGTC	1630
D <sub>b</sub>	LeuLeuTyrlYrsLysHisLysAlaTyrlYrsProIleGlyAsnCysProArgAsnThrVal	540
Q <sub>Y</sub>	AGAACRAAGCTGAGTGTCTTCTCAACCGTCAAAAACCGGTGTCTTCCCGGAAAC	1690
D <sub>b</sub>	LysGlyLysGlyLeuSerValLeuLeuSerHisAlaLysAlaProPhePheArgGlyAsp	560
Q <sub>Y</sub>	CAGAAAGAGGATCCGCTACTCAAAAACCAA	1720
D <sub>b</sub>	GlnGluLysAspProLeuLeuGlnAsnLys	570

## RESULT 4

Q8BVV9	PRELIMINARY;	PRT;	574 AA.
ID	Q8BVV9		
AC	Q8BVV9;		
DT	01-MAR-2003 (TREMBLrel. 23, Created)		
DT	01-MAR-2003 (T-EMBLrel. 23, Last sequence update)		
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)		
DE	Glycoprotein.		
GN	GNPB.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Skin;		
RX	MEDLINE=22354683; PubMed=12466851;		
RA	THE TANTOM Consortium,		
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;		
RT	"Analysis of the mouse transcriptome based on functional annotation of		
RT	60,770 full-length cDNAs."		
RL	Nature 420:563-573(2002).		
DR	EMBL: AK076347; BAC36306.1; -.		
DR	MGI: 1914765; Gpmdb.		
DR	GO: GO:000687; C:integral to plasma membrane; IDA.		
DR	GO: GO:0005194; F:cell adhesion molecule activity; IDA.		
DR	GO: GO:0008201; F:heparin binding; IDA.		
DR	GO: GO:0005178; F:integrin binding; IDA.		
DR	GO: GO:0007155; P:cell adhesion; IDA.		
DR	InterPro: IPR000583; GATase_2.		
DR	InterPro: IPR000601; PKD.		
DR	Pfam: PF00801; PKD; 1.		
DR	SMART: SM00089; PKD; 1.		
DR	PROSITE: PS00443; GATASE_TYPE_II; 1.		
DR	PROSITE: PS50093; PKD; 1.		
SQ	SEQUENCE 574 AA, 63675 MW; BGAE99DD5DB7C6D0 CRC64;		

US-10-039-272-1 (1-2661) x Q8BVV9 (1-574)

QY	60	ATGGAATGTCCTACTATTCTCTGGAAATTCCTGCTCTGCTGCAAGAAATTCGCACTTGAT	119
Db	1	MetGluserLeuCyGglyValLeuGlyPheLeuLeuLeuAlaGlyLeuProLeuGln	20
QY	120	GGCGGCAAAACGATTCATGATGTCTGGGCAATGAAGACCTTCGTCTACATCAGGGAG	179
Db	21	AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu	40

Qy 1211 GAGCTCTGATGCGGTCGATGCGCTCAAGAGCTCCCTAATAGAGACTTTGTGCTGACCTGC 1270  
D 401 AspValProMetProThrProGlnProAlaAsnSerLeuMetAspPheThrValThrCys 420  
Qy 1271 CAAGGGAGCATCCACCGAGGTCTGTACCATCATTTCTGACCCCACTGCGAGATCAC 1330  
D 421 LysGlyAlaThrProMetGluAlaCysThrIleSerAspProThrCysGlnIleAla 440  
Qy 1331 CAGAACACAGCTGTCAGCCCTGTGGATGTGTGATGATGTGTGCTGCTGCTGATGAGACGA 1390  
D 441 GlnAsnArgValCysSerProValAlaValAspGlyLeuCysLeuLeuSerValArgArg 460  
Qy 1391 ACCTTCATGCTGCGGACCTACTGTGTGACCTCACCTCGGGGATGACACAGCCTG 1450  
D 461 AlaPheAsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAlaSerLeu 480  
Qy 1451 GCTCTCAGCAGCACCTGATTTCTGCTCAGACAGACCCAGCCTCGCCTTTAAGGATG 1510  
D 481 AlaLeuThrSerThrLeuIleSerIleProGlyLysAspProAspSerProLeuArgAla 500  
Qy 1511 GCAAAACAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1570  
D 501 ValAsnGlyValLeuIleSerIleGlyCysLeuAlaValLeuValThrMetValThrIle 520  
Qy 1571 TTGGTGTACAAAACACAGAGGATACACACCAATAGAAATAGTCTGGGATGTGTC 1630  
D 521 LeuLeuTyrLysLysHisLysAlaTyrLysProIleGlyAsnCysProArgAsnThrVal 540  
Qy 1631 AGAAGCAAGGCTGAGTCTGCTTCTCAACCTGCAAAAGCCGTGCTTCTCCCGGAAAC 1690  
D 541 LysGlyLysGlyLeuSerValLeuLeuSerHisAlaLysAlaProPhePheArgGlyAsp 560  
Qy 1691 CAGGAAAGATCCGCTACTCAAAACCAA 1720  
D 561 GlnGluLysAspProLeuLeuGlnAspLys 570  
RESULT 5  
Q8BXL4 PRELIMINARY; PRT; 574 AA.  
ID Q8BXL4;  
AC Q8BXL4;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Glycoprotein.  
GN GPNMB.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Retina;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
60,770 full-length cDNAs."  
RL Nature 420:563-573(2002).  
DR EMBL; AK04764; BAC32074.1; -.  
DR MGD; MGI:1934765; Gpmb.  
DR GO; GO:0005887; C:integral to plasma membrane; IDA.  
DR GO; GO:0005194; P:cell adhesion molecule activity; IDA.  
DR GO; GO:0008201; F:heparin binding; IDA.  
DR GO; GO:0005178; F:integrin binding; IDA.  
DR GO; GO:0007155; P:cell adhesion; IDA.  
DR InterPro; IPR000583; GAtase\_2.  
DR InterPro; IPR000601; PKD.  
DR Pfam; PF00801; PKD; 1.  
DR SMART; SM00089; PKD; 1.  
DR PROSITE; PS00443; GATASE\_TYPE\_II; 1.  
DR PROSITE; PS00093; PKD; 1.  
SQ SEQUENCE 574 AA; 63693 MW; DC0D48485BA2E9D1 CRC64;

Alignment Scores: 7.5e-202 Length: 574  
Pred. No.: 2166.00 Matches: 401  
Score: 82.46% Conservative: 69  
Percent Similarity: 70.35% Mismatches: 84  
Best Local Similarity: 44.66% Indels: 17  
Query Match: 11 Gaps: 2  
DB: US-10-039-272-1 (1-2661) x Q8BXL4 (1-574)  
Qy 60 ATGGAATGCTCTACTATTTCTCTGGATTCTCTCTGGTCTGCAAGATTGCCACTTGAT 119  
D 1 MetGluSerLeuCysGlyValLeuGlyPheLeuLeuLeuAlaAlaGlyLeuProLeuGln 20  
Qy 120 GCCGCAAAACGATTTTCATGATGCTGGGCAATGAAGACCTTCTGCTTACATGAGGAG 179  
D 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAsnHisMetArgGlu 40  
Qy 180 CACAATCAATTAATGCTGCTTCTGATGAAATGCTGAATGCTGAATGAATAAATCTTACCCA 239  
D 41 HisAsnGlnLeuArgGlyTrpSerSerAspGluAsnGluTrpAspGluHisLeuTyrPro 60  
Qy 240 GTCTGGAAGCGGAGACATGAGGTGGAAACTCTCTGGAAGGAGGCGCTGTGAGGCG 299  
D 61 ValTrpArgArgGlyAspGlyArgTrpLysAspSerTrpGluGlyGlyArgValGlnAla 80  
Qy 300 GTCCTGACCAAGTACTCACCAGCCTCGTGGGCTCAAAATATAAATTTGCGGTGAACCTG 359  
D 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100  
Qy 360 ATATTCCCTAGATGCCAAAGAAAGATGCCATGCAACATAGTCTCTATGAGAGACTGC 419  
D 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120  
Qy 420 AGAAATGAGGCTGTTATCTCTCATCATATGTTTACAATGTCAGACATGCTGTCAGAG 479  
D 121 ArgAsnAspLeuGlyLeuThrSerAspLeuHisValTyrAsnTrpThrAlaGlyAlaAsp 140  
Qy 480 GACAGTGACGGGAAATGGCACCAGCCCAAGCCATCATAGCTCTTCCCTGATGGGAA 539  
D 141 AspGlyAspTrpGluAspGlyThrSerArgSerGlnHisLeuArgPheProAspArg 160  
Qy 540 CTTTTTCTCACCCCGGATGGAGACATGGAATTTTCATCTACCTCTTCCACACTT 599  
D 161 ProPheProArgProHisGlyTrpLysLysTrpSerPheValTyrValPheHisThrLeu 180  
Qy 600 GGTCACTATTCCAGAAATTTGGACGATGTTTCAGTCAGAGTCTTCTGTGAACACAGCAAT 659  
D 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerAlaArgValSerIleAsnThrValAsn 200  
Qy 660 GTGACACTTGGGCTCAACTCATGGAAGTGAAGTGTCTACAGAGACATGGACGGCATAT 719  
D 201 LeuThrAlaGlyProGlnValMetGluValThrValPheArgArgTyrGlyArgAlaTyr 220  
Qy 720 GTTCCCATCCCAAGTGAAGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 779  
D 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProValPheVal 240  
Qy 780 ACTATGTTCCAGAAAGACGATCAAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838  
D 241 ThrMetSerGlnLysAsnAspArgAsnLeuSerAspGluIlePheLeuArgAspLeuPro 260  
Qy 839 ATTATGTTGATGCTCTGATTCATGATTCATGATTCATGATTCATGATTCATGATTCAT 898  
D 261 IleValPheAspValLeuIleHisAspProSerHisPheLeuAsnAspSerAlaIleSer 280  
Qy 899 TACAAGTGGAGCTTCGGGGATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 958  
D 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300  
Qy 959 CACAGCATGTGCTCAATGAACCTTCAACCTTCAACCTTCAACCTTCAACCTTCAACCTT 1018  
D 301 HisThrTyrValLeuAsnGlyThrPheAsnLeuAsnLeuThrValGlnThrAlaValPro 320



## RESULT 7

Q8BVA0	PRELIMINARY;	PRT;	526 AA.
ID	Q8BVA0		
AC	Q8BVA0;		
DT	01-MAR-2003	(TREMBLrel. 23, Created)	
DT	01-MAR-2003	(TREMBLrel. 23, Last sequence update)	
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)	
DE	Glycoprotein.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Urinary bladder;		

Alignment Scores:			
Pred. No.:	3,91e-187	Length:	526
Score:	2015.00	Matches:	372
Percent Similarity:	83.17%	Conservative:	65
Best Local Similarity:	71.13%	Mismatches:	72
Query Match:	41.55%	Indels:	17
Pe.	11	Gaps:	2

Qy	60	ATGGAATGCTCTACTATTTCTCTGGGATTTCTGCTCTGGCTGCAAGATTGCCACTTGAT	119
Db	1	MetGluSerLeuGlyValLeuGlyPheLeuLeuLeuAlaLaglyLeuProLeuGln	20
Qy	120	GCCGCCAAACGATTTCATGTGCTCTGGCAATCAAAAGACCTTCTGCTTACATGAGGAG	179
Db	21	AlaAlaIysArgPheArgPheValLeuGlyHisGlnTyrProAspHisMetArgGlu	40
Qy	180	CACAAATCAATTAATGGCTGCTTCTGATGAAATGACTGGAATGAAAACTCTACCCA	239
Db	41	HisAsnGlnLeuArgGlyTyrSerSerAspGluAsnGluTrpAspGluHisLeuTyrPro	60
Qy	240	GTGTGGAAGCCGGAGACATGAGTGGNAAACTCTCTGGAAGGAGCGCGTGCAGCG	299
Db	61	ValTrpArgGlyAspGlyArgTrpIysAspSerTrpGluGlyGlyArgValGlnAla	80
Qy	300	GTCCTGCACGACTCACACGCCCTCTGTGGGCTCAAAATAATACATTTCCGGTGAACCTG	359
Db	81	ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu	100
Qy	360	ATATTCCCTAGATGCCAAAGAGATGCCAATGCCAACATAGTCTATGAGAAGAACTGC	419
Db	101	ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys	120
Qy	420	AGAAATGAGCGTGGTTTATCTGCTGATCCATATGTTTACAACTGGACACGATGTCAGAG	479
Db	121	ArgAsnAspLeuGlyLeuThrSerAspLeuHisValTyrAsnTrpThrAlaGlyAlaAsp	140
Qy	480	GACATGACGGGGAAAATGGACCGCCAAAGCCATCATACGTTCTCCCTGATGGGAA	539
Db	141	AspGlyAspTrpGluAspGlyThrSerArgSerGlnHisLeuArgPheProAspArg	160
Qy	540	CCTTTTCTCCACCCCGGATGGAAGAGATGGAATTTTCATCTAGCTCTCCACACACT	599
Db	161	ProPheProArgProHisGlyTrpIysIysTrpSerPheValTyrValPheHisThrLeu	180
Qy	600	GGTCAGTATTTCCAGAAATGGGACGATGTTTCAGTGAGAGTTTCTGTCAAACACGCCAAT	659
Db	181	GlyGlnTyrPheGlnLysLeuGlyArgCysSerAlaArgValSerIleAsnThrValAsn	200
Qy	660	GTGACACTTGGCCCTCAACTCATGGAAGTGACTGCTACAGAAGACATGGACGGCATAT	719
Db	201	LeuThrAlaGlyProGlnValMetGluValThrValPheArgArgTyrGlyArgAlaTyr	220
Qy	720	GTTCCCATCGCAAGATGATGTACGTGGTAAACAGATCAGATTCTCTGTGTTGTG	779
Db	221	IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProValPheVal	240



```
QY 780 ACTATGTCAGAGAACGATCGAAATTCATCCGAGAACCTTCCC-AAAGATCTCCCC 838
Db |||||
QY 241 ThrMetSer:GlnLysAsnAspArgAsnLeuSerAspGluLeuPheLeuArgAspLeuPro 260
Db |||||
QY 839 ATTATGTTGATGCTCTGATTCATGATCCTAGCCACTTCTCAATATTCTACCAATAAC 898
Db |||||
QY 261 IleValPheAspValLeuIleHisAspProSerHisPheLeuAsnAspSerAlaIleSer 280
Db |||||
QY 899 TACAGTGGAGCTTCGGGATATATACATGGCTGTTTGTTCACCAATCATATCTGCAAT 958
Db |||||
QY 281 TyrLysTrpAsnPhedGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300
Db |||||
QY 959 CACAGCTATGCTCAATCGAAGCTTACAGCTTAACTCAGCTGCTGAAAGCTGCAGACCA 1018
Db |||||
QY 301 HisThrTyrValLeuAsnGlyThrPheAsnLeuAsnLeuThrValGlnThrAlaValPro 320
Db |||||
QY 1019 GGACCTTGTCGG-----CCACGGCCACCA 1042
Db |||||
QY 321 GlyProCysProProSerProSerThrProProProProSerThrProProSerPro 340
Db |||||
QY 1043 CCACCCAGACCT-----TCAAAACCCACCCCTTTTAGGACCTGCTGGTAC 1090
Db |||||
QY 341 ProProSerProLeuProThrLeuSerThrProSerProSerLeuMetProThrGlyTyr 360
Db |||||
QY 1091 AACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCCAGATTAAACAGATATGCCAC 1150
Db |||||
QY 361 LysSerMetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyr 380
Db |||||
QY 1151 TTCAAGCACCATCACAATTGTAGAGGAATCTTAGAGGTTAATCATCATCCAGATGACA 1210
Db |||||
QY 381 PheArgAlaThrIleThrIleValGluGlyIleLeuGluValSerIleMetGlnIleAla 400
Db |||||
QY 1211 GAGCTCCTGATGCGCGTCCGCTGAGAGCTCCCTAATAGACTTTGTGCTGACCTGC 1270
Db |||||
QY 401 AspValProMetProThrProGlnProAlaAsnSerLeuMetAspPheThrValThrCys 420
Db |||||
QY 1271 CAAGGAGCATTCACCCAGAGCTGTGTACCATCATTTCTGACCCACCTGCAGATCAC 1330
Db |||||
QY 421 LysGlyAlaThrProMetGluAlaCysThrIleIleSerAspProThrCysGlnIleAla 440
Db |||||
QY 1331 CAGAACACAGCTGAGCCCTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1390
Db |||||
QY 441 GlnAsnArgValCysSerProValAlaValAspGlyLeuCysLeuLeuSerValArgArg 460
Db |||||
QY 1391 ACCTTCATGAGTCTGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1450
Db |||||
QY 461 AlaPheAsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAlaSerLeu 480
Db |||||
QY 1451 GCTCTCAGAGCACCCCTGATTTCTGTTCTCCTGACAGAGACCCAGCTGCGCTTTAAGGATG 1510
Db |||||
QY 481 AlaLeuThrSerThrLeuIleSerIleProGlyLysAspProAspSerProLeuArgAla 500
Db |||||
QY 1511 GCACAGAGTCCCTGATCTCGTGGTGTGCTGGCCATATTGTCACTGTGATCTCCCTC 1570
Db |||||
QY 501 ValAsnGlyValLeuIleSerIleGlyCysLeuAlaValLeuValThrMetValThrIle 520
Db |||||
QY 1571 TTGGTGTAC 1579
Db |||||
QY 521 LeuLeuTyr 523
```

## RESULT 8

```
Q96F58
ID Q96F58 PRELIMINARY; PRT; 206 AA.
AC Q96F58;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Similar to glycoprotein (Transmembrane) nmb (Hypothetical protein
DE GPNMB).
GN GPNMB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99063792; PubMed=9847074;
RA Sulston J.E., Waterston R.;
RT "Toward a complete human genome sequence.";
RL Genome Res. 8:1097-1108(1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Kalicki J., Kramer J.;
RT "The sequence of Homo sapiens BAC clone CTA-271G13.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Waterston R.H.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
RN [9]
RP SEQUENCE FROM N.A.
RA Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,
RA Phelan M., Farmer A.;
RT "Cloning of human full-length CDSs in BD Creator(TM) System Donor
RT vector.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC011595; AAH11595.1; -.
DR EMBL; AC005082; AAP22337.1; -.
DR EMBL; BT007074; AAP35737.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
KW Hypothetical protein; Transmembrane.
SQ SEQUENCE 206 AA; 23945 MW; 30CBDE6928D73FBD CRC64;
```

## Alignment Scores:

```
Pred. No.: 4,26e-91 Length: 206
Score: 1029.00 Matches: 182
Percent Similarity: 98.38% Conservative: 0
Best Local Similarity: 98.38% Mismatches: 3
Query Match: 21.22% Indels: 0
DB: 4 Gaps: 0
```

US-10-039-272-1 (1-2661) x Q96F58 (1-206)

```
QY 60 ATGGAATGCTCTACTATTCTCTGGGATTTCTGCTCTGCTGCAAGATTGCCACTTGAT 119
Db |||||
QY 1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20
Db |||||
QY 120 GCCGCCAACGATTTTCATGATGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGAG 179
Db |||||
QY 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgL 40
Db |||||
QY 180 CACAATCAATTAATGCTGCTTCTGATGAAATGACTGGATGAAATCTTACCCA 239
Db |||||
```

Db	41	HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro	60
QY	240	GTGTGGAGCGGGAGAGCATGAGGTGGAAAACTCTGGAGGGAGCGCGTGTGAGCGC	299
Db	61	ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla	80
QY	300	GTCCTGACAGTGAATCACCAGCCCTGTGGGGCTCAATATAACATTTGCCGTGAACCTG	359
Db	81	ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu	100
QY	360	ATATTCCTAGATGCCAAAAAGAGATGCCAATGGCAACATAGTCTATGAGAGAACTGC	419
Db	101	IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTrpGluLysAsnCys	120
QY	420	AGAAATGAGCGTGGTTTATCTGCTGATCCATATGTTTCAACTGACAGCATGTCAGAG	479
Db	121	ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu	140
QY	480	GACAGTACGGGGAAAAATGGCACCGGCCAAAGCCATCAACGCTTCTCCCTGATGGAAA	539
Db	141	AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys	160
QY	540	CTTTTCTCCACACCCCGGATGGAGATGGAATTTTCATCTAGCTTCTCCACACACTT	599
Db	161	ProPheProHisHisProGlyTrpArgTrpAsnPheIleTyrValPheHisThrLeu	180
QY	600	GGTCAGTATTTCCAG 614	
Db	181	AspTrpLeuLeuGln 185	
RESULT 10			
QY	Q9CZB2	PRELIMINARY; PRT; 626 AA.	
AC	Q9CZB2		
DT	01-JUN-2001	(TrEMBLrel. 17, Created)	
DT	01-JUN-2001	(TrEMBLrel. 17, Last sequence update)	
DT	01-JUN-2003	(TrEMBLrel. 24, Last annotation update)	
DE	N/A.		
GN	SI OR SI		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Embryo;		
RX	MEDLINE=21085660; PubMed=11217851;		
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,		
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,		
RA	Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,		
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,		
RA	Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,		
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,		
RA	Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,		
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,		
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarts P.,		
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,		
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,		
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,		
RA	Hayashizaki Y.;		
RT	"Functional annotation of a full-length mouse cDNA collection.";		
RL	Nature 409:685-690(2001).		
DR	EMBL; AK012808; BAB28486.1; -		
DR	MGI; MGI:98301; Si.		
DR	InterPro; IPR000601; PKD.		
DR	Pfam; PF00801; PKD; 1		
DR	SMART; SM00089; PKD; 1.		
DR	PROSITE; PS50093; PKD; 1.		

Db	41	HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro	60
QY	240	GTGTGGAGCGGGAGAGCATGAGGTGGAAAACTCTGGAGGGAGCGCGTGTGAGCGC	299
Db	61	ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla	80
QY	300	GTCCTGACAGTGAATCACCAGCCCTGTGGGGCTCAATATAACATTTGCCGTGAACCTG	359
Db	81	ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu	100
QY	360	ATATTCCTAGATGCCAAAAAGAGATGCCAATGGCAACATAGTCTATGAGAGAACTGC	419
Db	101	IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTrpGluLysAsnCys	120
QY	420	AGAAATGAGCGTGGTTTATCTGCTGATCCATATGTTTCAACTGACAGCATGTCAGAG	479
Db	121	ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu	140
QY	480	GACAGTACGGGGAAAAATGGCACCGGCCAAAGCCATCAACGCTTCTCCCTGATGGAAA	539
Db	141	AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys	160
QY	540	CTTTTCTCCACACCCCGGATGGAGATGGAATTTTCATCTAGCTTCTCCACACACTT	599
Db	161	ProPheProHisHisProGlyTrpArgTrpAsnPheIleTyrValPheHisThrLeu	180
QY	600	GGTCAGTATTTCCAG 614	
Db	181	GlyTrpLeuLeuGln 185	
RESULT 9			
QY	Q8IXJ5	PRELIMINARY; PRT; 206 AA.	
AC	Q8IXJ5		
DT	01-WAR-2003	(TrEMBLrel. 23, Created)	
DT	01-WAR-2003	(TrEMBLrel. 23, Last sequence update)	
DT	01-WAR-2003	(TrEMBLrel. 23, Last annotation update)	
DE	Glycoprotein mmb-like protein.		
GN	GNMB.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Lennarz V.;		
RT	"Identification and Characterization of T cell defined tumor antigens		
RL	in the melanoma model M27.";		
RL	Thesis (2002), Department of Biological Sciences,		
RL	Johannes Gutenberg-University of Mainz, Mainz, Germany.		
DR	EMBL; AJ505015; CAD43718.1; -		
SQ	SEQUENCE 206 AA; 24003 MW; 30C8E5A28D73FBD CRC64;		
Alignment Scores:			
Pred. No.:	2,05e-90	Length:	206
Score:	1022.00	Matches:	181
Percent Similarity:	97.84%	Conservative:	0
Best Local Similarity:	97.84%	Mismatches:	4
Query Match:	21.07%	Indels:	0
DB:	4	Gaps:	0
US-10-039-272-1 (1-2661) x Q8IXJ5 (1-206)			
QY	60	ATGGAATGCTCTACTATTCTCTGGGATTTCTGCTGGTGGCAAGATGCCACTTGAT	119
Db	1	MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp	20
QY	120	GCCGCCAAACGATTTGATGCTGCTGGCAATCAAGACCTTCTGCTTACATGAGGAG	179
Db	21	AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu	40
QY	180	CACAAATCAATAAATGGTGGTCTTCTGATGAAATGACTGGAATGAAAACTCTACCA	239

SQ SEQUENCE 626 AA; 66301 MW; 7EC0A06C63212674 CRC64;

Alignment Scores:		
Pred. No.:	2,138-53	Length:
Score:	644.50	Matches:
Percent Similarity:	40.94%	Conservative:
Best Local Similarity:	27.25%	Mismatches:
Query Match:	13.29%	Indels:
DB:	11	Gaps:
		21
		636

US-10-039-272-1 (1-2661) x Q9CZB2 (1-626)

78	TTCCTGGGATTTGTCTCTCGCTCAAGATTGCCACTTGATGCC-----GCCAAACGA	131
QY	:::     :::     :::     :::     :::     :::     :::     :::	
Db	PheLeuProValLeuValLeuSerAlaLeuLeuAlaValGlyAlaLeuGluGlySerArg	27
QY	132 TTTTCATGATGCTGGGGCAATGAAAGACCTTCTGCTTACATCAGGAGCACAAATCAATTA	191
Db	:::     :::     :::     :::     :::     :::     :::	
Db	28 AsnGlnAspTrpLeuGlyValProArgGlnLeuVal-----	39
QY	192 AATGCCTCGTCTTCGTATGATAAATGACTTGGATGAAAACCTACCAGTGTGGAGCGG	251
Db	:::     :::     :::     :::     :::     :::     :::	
Db	40 -----ThrLysThrTrpAsnArgGlnLeuTyProGluTrpThr---	52
QY	252 GGGACATGAGGTGGAAAAAATCCTCTGGAAGGAGCCGTGTGACGGCTCTGACCACT	311
Db	:::     :::     :::     :::     :::     :::     :::	
Db	53 ---GluValGlnGlySerAsnCysTrpArgGlyGlyGlnValSerLeuArgValIleAsn	71
QY	312 GACTCACCGACCTCGTGGGCTCAAATATACATTTGCGGTGAACCTGATATTCCTTAGA	371
Db	:::     :::     :::     :::     :::     :::     :::	
Db	72 AspGlyProThrLeuValGlyAlaAsnAlaSerPheSerIleAlaLeuHisPheProGly	91
QY	372 TGCCAAAAGGAAGATGCCATGGGCAACAATGCTCTATGAGAAGAAC--TGCAGAATAGAG	428
Db	:::     :::     :::     :::     :::     :::     :::	
Db	92 SerGlnLysValLeuProAspGlyGlnValIleTrpAlaAsnAsnThrIleIleAsnGly	111
QY	429 CTGTGGTTTTCTGTGTATCCCATATGTTTACAACTGGACAGCATGGTCAGAGACAGTAGAC	488
Db	:::     :::     :::     :::     :::     :::     :::	
Db	112 SerGlnValTrpGlyGlyGlnProValTyPro-----GlnGluProAspAsp	127
QY	489 GGGGAAAATGGCACCGGCCAAAGCCATCAATAGCTCTTCCCTGATGGGAAAACCTTTTCT	548
Db	:::     :::     :::     :::     :::     :::     :::	
Db	128 Ala-----CysValPheProAspGlyGlyProCysPro	138
QY	549 CACACACCCCGGATGGAGAAGATGGAATTTTCATCTACGTCCTCCACACACTTGGTCAGAT	608
Db	:::     :::     :::     :::     :::     :::     :::	
Db	139 SerGlyProLysProProLysArgSerPheValTyValTrpLysThrTrpGlyLysTy	158
QY	609 TTCAGAAATGGGACGATGTTCACTGAGAGTTTCTGTGACACAGACCAATGTGCACATT	668
Db	:::     :::     :::     :::     :::     :::     :::	
Db	159 TrpGlnValLeuGlyGlyProValSerArgLeuSerIleAlaThrGlyHisAlaLysLeu	178
QY	669 GGSCCTCAACTCATGGAGTCACTGCTCTACAGAAGACATGGA---CGGGCATATGTTCCC	725
Db	:::     :::     :::     :::     :::     :::     :::	
Db	179 GlyThrHisThrMetGluValThrValTyHisArgArgGlySerGlnSerTyValPro	198
QY	726 ATCCGCAACAGTGAAGAATGTGCTCGTGAACACATCAGATTCTCTGTTGTGCACATG	785
Db	:::     :::     :::     :::     :::     :::     :::	
Db	199 LeuAlaHisAlaSerSerThrPheThrIleThrAspGlnValPropheSerValSerVal	218
QY	786 TTCCAGAA-----GAACGATCGAAATTCATCCGACGAAACCTTCCCCAAA	829
Db	:::     :::     :::     :::     :::     :::     :::	
Db	219 SerGlnLeuGlnAlaLeuAspGlyGluThrLysHisPheLeuArgAsnHis-----	235
QY	830 GATCTCCCCATTATGTTTGATGCTCTGATTCATGATCTTAGCCACTTCTCTCAATATTCT	889
Db	:::     :::     :::     :::     :::     :::     :::	
Db	236 ----ProLeuIlePheAlaLeuGlnLeuHisAspProSerGlyTyLeuAlaGluAla	253
QY	890 ACCATTAACTACAAGTGGAGCTTCGGGGATAATACTGGCCCTGTTGTTGTTCCACCAATCAT	949
Db	:::     :::     :::     :::     :::     :::     :::	
Db	254 AspLeuSerTyThrTrpAspPheGlyAspGlyThrGlyThrLeuIleIleSerArgAlaLeu	273
QY	950 ACTGTGAATCACACGATGTGCTCAATGGAAACCTTCATCGCCTTAACCTCACTGTGAAAGCT	1009

Db	274	AspValThrHisThrTyrLeuGluSerGlySerValThrAlaGlnValValLeuGlnAla	293
QY	1010	GCA-----	1012
Db	294	AlaIleProLeuValSerCysGlySerSerProValProGlyThrThrAspGlyTyrMet	313
QY	1012	-----	1012
Db	314	ProThrAlaGluAlaProGlyThrThrSerArgGlnGlyThrThrLysValValGly	333
QY	1013	---GCACGAGACCTTGTCCGCCACCGCCACCA	1042
Db	334	ThrThrProGlyGlnMetProThrThrClnProSerGlyThrThrValValGlnMetPro	353
QY	1042	-----	1042
Db	354	ThrThrGluValThrAlaThrThrSerSerGluGlnMetLeuThrSerAlaValIleAspThr	373
QY	1043	-----	1043
Db	374	ThrLeuAlaGluValSerThrThrGluGlyThrGlyThrThrProThrArgProSerGly	393
QY	1058	-----	1058
Db	394	ThrThrValAlaGlnAlaThrThrGluGlyProAspAla---SerProLeuLeuPro	412
QY	1100	-----	1123
Db	413	ThrGlnSerSerThrGlySerIleSerProLeuLeuAspThrAspThrIleMetLeu	432
QY	1124	-----	1162
Db	433	VallysArgGlnValProLeuAspCysValLeuTyrArgTyrGlySerPheSerLeuAla	452
QY	1163	ATCACAAATTGAGAGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGCTCGATG	1222
Db	453	LeuAspIleValGlnGly-----	466
QY	1223	CCGGTGGCATGGCTGAAAGCTCCCTAATAGACTTTGTGCTGACCTGCCAGGGAGCAATT	1282
Db	467	AlaValProPheSerGluGlyAspAlaPheGluLeuThrValSerCysGlnGlyGlyLeu	486
QY	1283	CCACGAGAGTCTGTACCACTTTCTGACCCACCTGCGAGATCACCAGACACAGTC	1342
Db	487	ProLysGluAlaCysMetAspIleSerSerProGlyCysGlnProProAlaGlnArgLeu	506
QY	1343	TGCAGCCCTGTGATGTGGATGATGTCTGCTGACTGTGAGAGCAACCTTC---AAT	1399
Db	507	CysGlnSerValProProSerProAspCysGlnLeuValLeuHisGlnValLeuLysGly	526
QY	1400	GGGTCTGGGACGTACTGTGTGAACCTCACCTGGGGATGACACAGCTGGCTCTCAG	1459
Db	527	GlySerGlyThrTyrCysLeuAsnValSerLeuAlaAspAlaAsnSerLeuAlaValAla	546
QY	1460	AGACCCCTGATTCTGTCTCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAACACT	1519
Db	547	SerThrGlnLeuValValProGlyGlnAspGlyGly-----LeuGlyGln	561
QY	1520	GCCTGATCTCCGTTGGCTGCTTGGCCATATTGTGCATCTGTATCTCCCTCTGTGTAC	1579
Db	562	AlaProLeuValGlyIleLeuLeuValLeuValAlaValAlaValLeuAlaSerLeuIle	581
QY	1580	AAAAACAC-----	1624
Db	582	HisArgHisArgLeuLysGlnGlySerValSerGlnMetProHisGlySerThrHis	601
QY	1625	-----	1666
Db	602	TrpLeuArgLeuProProValPheArgAlaAaGlyLeuGly-----	615
QY	1667	AAAGCCGTGTTCTCCCGGAAACCCAGGAAAGGATCCGCTACTCAAAACCAAGAA	1723

Db 616 -----GluAsnSerProLeuLeuSerGlyGlnGln 625

RESULT 11

O97884 PRELIMINARY; PRT; 461 AA.

AC O97884; 1999 (TREMELrel. 10, Created)

DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)

DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)

DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)

DE Melanocyte protein 17 (Fragment).

GN PMEL17.

OS Equus caballus (Horse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

OX NCBI\_TaxID=9796;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Skin;

RA Rieder S., Checa-Cortes M.L., Joerg H., Stranzinger G.;

RT "An equine sequence homologous to melanocyte protein 17 (PMEL17)

RL submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF076780; AAC97108.1; --

DR InterPro; IPR000601; PKD.

DR Pfam; PF00801; PKD; 1.

DR SMART; SM00089; PKD; 1.

DR PROSITE; PSS0093; PKD; 1.

FT NON\_TER 1

FT NON\_TER 461

SQ SEQUENCE 461 AA; 49334 MW; 12752AF6C1EC373D CRC64;

Alignment Scores:

Pred. No.: 7.45e-27 Length: 461

Score: 372.00 Matches: 101

Percent Similarity: 37.84% Conservative: 64

Best Local Similarity: 23.17% Mismatches: 119

Query Match: 7.67% Indels: 153

DB: 6 Gaps: 9

US-10-039-272-1 (1-2661) x O97884 (1-461)

QY 702 AGACATGGAGCGGCATATGTCCTCCATCGCAGCAAGTGAAGATGTGTACGTGTAACAGAT 761

Db 1 ArgGlySerGlnThrThrMetProLeuAlaHisSerArgSerAlaPheThrThrAsp 20

QY 762 CAGATTTCCTGTTGTGACTATGTTCCAGAGAACGATCGAAATTCATCCGACGAAACC 821

Db 21 GlnValProPheSerValSerValSerGlnLeuGlnAlaLeuAspGlyArgAsnLysHis 40

QY 822 TTCCC-AAAGATCCCTCCATTATGTTGATGCTGCTGATTCATGCTAGCCACTTCCTC 880

Db 41 PheLeuLysAsnGlnProLeuThrPheAlaLeuArgLeuHisAspProSerGlyThrLeu 60

QY 881 AATTATTCTACCAATTAACTACAAAGTGGAGCTTCGGGGATAAATACTGGCTTTGTTC 940

Db 61 AlaGlyAlaAspLeuSerThrThrAspPheGlyAspSerThrGlyThrLeuIleSer 80

QY 941 ACCAATCATCTGTGAATCACAGTATGCTCAATGGAACCTTCAGCCTTAACCTCACT 1000

Db 81 ArgAlaLeuValValThrHisThrThrLeuGluSerGlyProValThrAlaGlnValVal 100

QY 1001 GTGAAGCTGCACACCA-----GGACCTTGTCGCCACCG----- 1036

Db 101 LeuGlnAlaAlaIleProLeuThrSerCysGlySerSerProValProGlyThrThrGly 120

QY 1036 ----- 1036

Db 121 GlyThrValProThrAlaGluAlaProGlyThrThrAlaGlyGlnValProThrAlaAsp 140

QY 1037 -----CCACCACCCAGACCTTCACAAACCCACC----- 1066

Db 141 ValValAsnThrThrProGlyGlnValProThrAlaGluProSerArgThrThrAlaVal 160

QY 1067 -----CCTCTTTTAGGACCT 1081

Db 161 GlnValProThrThrGluValIleSerThrThrProValGlnValProThrAlaGluAsp 180

QY 1082 GCTGGTGACAAACCCCTGGAGCTGAGTAGG----- 1111

Db 181 IleGlyThrThrProGluGlnValSerThrProGluSerLeuGlyThrThrLeuAlaGlu 200

QY 1112 -----ATTCTCATGAA----- 1123

Db 201 MetProLeuGlnArgLeuLysGlyIleProGluValSerAsnSerArgSerLeuLeu 220

QY 1123 ----- 1123

Db 221 GlyHisSerAlaGlnValThrValGlnSerTrpTrpLysProGlnTrpArgGlyThrGln 240

QY 1123 ----- 1123

Db 241 ProLeuSerLeuArgValGlnMetProAlaHisSerCysTyrArgArgAsnTyrArgSer 260

QY 1123 ----- 1123

Db 261 GlnSerProLeuLeuAspGlyThrAlaThrLeuPheLeuValLysArgGlnValProLeu 280

QY 1124 AACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATCACAAATTGTAGAGGAATC 1183

Db 281 AspCysValLeuTyrArgTyrGlySerPheSerLeuThrLeuAspIleValGlnGly--- 299

QY 1184 TTAGAGGTTAATCATCATCCAGATGACAGACGTCCTGATGCCGTCCTGCGCTGAAAGC 1243

Db 300 -----IleGluSerAlaGluIleLeuGlnAlaValProSerSerGluGly 314

QY 1244 TCCTTAATAGACTTTGCTGACCTGCCAAGGAGCATTTCCACGAGGCTGTGTACCATC 1303

Db 315 AspAlaPheGluLeuThrValSerCysGlnGlyGlyLeuProLysGluAlaCysMetAsp 334

QY 1304 ATTTCTGACCCACCTCGAGATACCCAGAACACAGCTGTGAGCCTGTGATGGATG 1363

Db 335 IleSerSerProGlyCysGlnProProAlaGlnArgLeuCysGlnProValThrProAsn 354

QY 1364 GAGATGTGCTGCTGACTGTGAGACGAACTTC---AATGGGTCTGGGACGTACTGTGTG 1420

Db 355 ProAlaCysGlnLeuValLeuHisGlnValLeuLysSerGlySerGlyThrThrCysLeu 374

QY 1421 AACCTCACCTCGGGGATGACAAAGCCTGCTCTCACGACGACCCCTGATTTCTGTCT 1480

Db 375 AsnValSerLeuAlaAspAlaAsnSerLeuAlaMetValSerThrGlnLeuValMetPro 394

QY 1481 GACAGAGACCCAGCCTCGCCTTTAAGGATGGCAACAGTCGCTGATCTCCGTTGGCTGC 1540

Db 395 GlyGlnGluAlaGly-----LeuGlyGlnAlaProLeuPheValGlyIle 409

QY 1541 TTGGCCATATTGTCTACT---GTGATCTCCCTCTGGTGTACAAAAA 1585

Db 410 LeuLeuValLeuMetAlaMetValLeuAlaSerLeuIleThrArgArg 425

RESULT 12

O93391 PRELIMINARY; PRT; 270 AA.

AC O93391; 1998 (TREMELrel. 08, Created)

DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)

DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)

DE Melanosomal matrix protein (Fragment).

GN MP115.

OS Coturnix coturnix (Common quail).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Coturnix.

OX NCBI\_TaxID=9091;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99119493; PubMed=9918878;

RA Oberst C., Hartl M., Weiskirchen R., Bister K.;  
 RT "Conditional cell transformation by doxycycline-controlled expression  
 of the MC29 v-myc allele.";  
 RL Virology 253:193-207(1999).  
 DR EMBL; AF077328; AAD12180.1; -.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 KW Matrix protein.  
 FT NON TER 1  
 SQ SEQUENCE 270 AA; 26785 MW; 564AA46F0B875E63 CRC64;  
  
 Alignment Scores:  
 Pred. No.: 3.84e-17 Length: 270  
 Score: 271.50 Matches: 71  
 Percent Similarity: 48.03% Conservative: 39  
 Best Local Similarity: 31.00% Mismatches: 94  
 Query Match: 5.60% Indels: 25  
 DB: 13 Gaps: 6  
  
 US-10-039-272-1 (1-2661) x O93391 (1-270)  
 QY 1067 CCTTCTTAGACCTGTGGTGACACCCCTGGAGCTAGTAGTCTCGATGAAAC 1126  
 DB 50 ProThrAlaGlySerThrAlaGluProLeuLeuValLysArgGlnAlaProGluSer 69  
 QY 1127 -----TGCCAGATTAAAGATATGCGCATTTCAGGCCACCATCACAAATGTA 1174  
 DB 70 GluProThrGlyCysValLeuTyArgTyrGlySerPheSerThrGluLeuAsnIleVal 89  
 QY 1175 GAGGAAATCTTAGAGTTAAATCATCATCCAGATGACAGACGTCCTGATGCCGCTCCATGG 1234  
 DB 90 GlnGlyIleGluAsnValAlaIleValGln-----ValValProAlaAlaProGlu 106  
 QY 1235 CCTGAAGCTCCCTAATAGACTTTGTGTGACCTGCCAGGAGGAGCATTCACAGAGTC 1294  
 DB 107 GlySerGlyAsnSerValGluLeuThrValThrCysGluGlySerLeuProGluGluVal 126  
 QY 1295 TGTACCATCATTTCTGACCCCTGCGAGATCACCAGAACACAGTCGACGCTGTCG 1354  
 DB 127 CysThrValValAlaAspAlaGluCysArgThrAlaGlnMetGlnThrCysSerAlaVal 146  
 QY 1355 GATGTGGATGAGTGTCTGTGACTGTGAGAGCAACCTTCAATGGTCTGGAGCCTAC 1414  
 DB 147 AlaProAlaProGlyCysGlnLeuValLeuArgGlnAspPheAsnGlnSerGlyLeuTyr 166  
 QY 1415 TGTGTGAACCTCACCTCGGGGATGACACAGCTGCTCTCAGAGACCTGATTTCT 1474  
 DB 167 CysLeuAsnValSerLeuAlaAsnGlyAsnGlyLeuAlaValAlaSerThrArgValAla 186  
 QY 1475 GTTCCTGACAGACCCAGCTCGCTTTAAGGATGCAAAACAGTCGCTGATCTCCGTT 1534  
 DB 187 ValGlyGlyAlaSerProAla-----AlaGlyGlyThrThrLeuThrVal 201  
 QY 1535 GGCTGTGGCCATTTGTGCTGATCTCCCTCTTGTGTACAAAAACACAGAA 1594  
 DB 202 GlyLeuLeuLeuIleValAlaAlaLeuGlyThrAlaAlaTyrThrTyrArgGValLys 221  
 QY 1595 TACACCCATAGAAATAGTCTGGGAATGTGTGAGA----- 1633  
 DB 222 TyrSerProLeuLeuProThrAlaProGlnValSerArgProHisSerTrpLeuProPro 241  
 QY 1634 AGCAAGGCTGAGTGTCTTCTCAACCGTGCAAAAGCCGTGTTCTTCCCGGA----- 1687  
 DB 242 GlyAlaAlaLeuArgLeuLeuLeuArgGlnAla-----PheGlyGlyAlaPro 257  
 QY 1688 AACGAGAAAGGATCCGCTACTCAA 1714  
 DB 258 SerGlyGluSerProLeuLeuArg 266  
  
 RESULT 13  
 QYQY67 PRELIMINARY; PRT; 236 AA.  
 AC QYQY67;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)  
 DE Mutant matrix melanosomal protein silver (Fragment).  
 GN SI.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20063689; PubMed=10594241;  
 RA Martinez-Esparza M., Jimenez-Cervantes C., Bennet D.C., Lozano J.A.,  
 RA Solano F., Garcia-Borron J.C.;  
 RT "The mouse silver locus encodes a single transcript truncated by the  
 RT silver mutation.";  
 RL Mamm. Genome 10:1168-1171(1999).  
 DR EMBL; AF119092; AAF19761.1; -.  
 DR MGD; MGI:98301; Si  
 DR InterPro; IPR007110; Ig-like.  
 FT NON TER 1  
 SQ SEQUENCE 236 AA; 24293 MW; 3ECB812D4DF10EE7 CRC64;  
  
 Alignment Scores:  
 Pred. No.: 7.98e-15 Length: 236  
 Score: 247.50 Matches: 62  
 Percent Similarity: 47.69% Conservative: 31  
 Best Local Similarity: 31.79% Mismatches: 85  
 Query Match: 5.10% Indels: 17  
 DB: 11 Gaps: 5  
  
 US-10-039-272-1 (1-2661) x Q9QY67 (1-236)  
 QY 1019 GGACCTTGTGCGCCACCCACCCAGACCTTCAAAACCCACCCCTCTTTAGGA 1078  
 DB 39 GlyProAspAlaSerProLeuLeuProThrGlnSerSer-----ThrGlySerIleSer 56  
 QY 1079 CTGCTGTGTGACAAACCC-----CTGGAGCTAGTAGATTCTGTATGAAACTGC 1129  
 DB 57 ProLeuLeuAspThrAspThrIleMetLeuValLysArgGlnValProLeuAspCys 76  
 QY 1130 CAGATTAAACAGATATGCCACTTTCAGCCACCATCATCAATTGTAGAGGAATCTTAGAG 1189  
 DB 77 ValLeuTyrArgTyrGlySerPheSerLeuAlaLeuAspIleValGlnGly----- 93  
 QY 1190 GTTAACATCATCCAGATGACAGACGTCCTGATCGGTGCCATGCGCTGAAAGCTCCCTA 1249  
 DB 94 -----IleGluSerAlaGluIleLeuGlnAlaValProSerSerGluGlyAspAla 110  
 QY 1250 ATAGACTTGTGCTGACCTGCCAAGGAGGAGCATTCACACGAGGCTCTACCATTTCT 1309  
 DB 111 PheGluLeuThrValSerCysGlnGlyGlyLeuProLysGluAlaCysMetAspIleSer 130  
 QY 1310 GACCCACCTCGCAGATCACCCAGAACACAGTCGTGAGCCCTGTGGATGTGGATGAGATG 1369  
 DB 131 SerProGlyCysGlnProProAlaGlnArgLeuCysGlnSerValProProSerProAsp 150  
 QY 1370 TGTCTGTGCTGCTGAGACGAACTTC-----AATGGGTCTGGGACGTACTGTGAACTC 1426  
 DB 151 CysGlnLeuValLeuHisGlnValLeuLysGlyGlySerGlyThrTyrCysLeuAsnVal 170  
 QY 1427 ACCCTGGGGATGACACAGACCTGCTCTCAGAGCACCCCTGATTCTGTCTCTGACAGA 1486  
 DB 171 SerLeuAlaAspAlaAsnSerLeuAlaValAlaSerThrGlnLeuValValProGlyGln 190  
 QY 1487 GACCCAGCTCGCTTTAAGGATGGCAACAGTCGCTGATCTCCCTGCTGCTGCTGGCC 1546  
 DB 191 AspGlyGly-----LeuGlyGlnAlaProLeuLeuValGlyIleLeuLeu 205  
 QY 1547 ATATTTGTCTGCTGATCTCCCTCTTGGTGTACAAAAACACAAAG 1591  
 DB 206 ValLeuValAlaValValLeuAlaSerLeuLeuIleHisArgHisArg 220  
  
 RESULT 14



```

DE Hypothetical protein.
GN DKFZP761H2017.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Amysdala;
RA Ansoerge M., Wirkner U., Mewes H.W., Weil B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL834349; CAD39014.1; -.
DR InterPro; IPR00601; PKD.
DR PROSITE; PS50093; PKD; 1.
KW Hypothetical protein.
SQ SEQUENCE 435 AA; 48329 MW; 737509A3F8263A4F CRC64;

Alignment Scores:
Pred. No.: 6.17e-10 Length: 435
Score: 198.50 Matches: 112
Percent Similarity: 37.06% Conservative: 67
Best Local Similarity: 23.19% Mismatches: 162
Query Match: 4.09% Indels: 142
DB: 4 Gaps: 27

US-10-039-272-1 (1-2661) x Q8N3G9 (1-435)
QY 240 GTGTGGAGCGG-----GGAGACATGAGGTGGAAAC-----TCTGG----- 278
Db 5 ValTrpSerArgLeuGlyArgIleLeuTrpLeuAlaCysLeuLeuProTrpAlaProAla 24
QY 279 -----AAGGAGCGCGTGCAGCGGTCTGTGACAGTACTGACAGCGCTCTGTCGGC 332
Db 25 GlyValAlaAlaGlyLeuTyrGluLeuAsnLeuThrThrAspSerProAlaThrGly 44
QY 333 TCAAAATAACATTTGGCGGTGAACCTGATATTCCTAGATGCCAAAGGAAGATGCCAAT 392
Db 45 AlaValValThrIleSerAlaSerLeuVal-----AlaLysAspAsn 58
QY 393 GGCACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGTGATCCATAT 452
Db 59 GlySerLeu-----AlaLeuProAlaAspAlaHis 68
QY 453 GTTTAC-----AACTGGACAGATGGTGCAGACAGTACGCGGGAATGGCACCGC 506
Db 69 LeuTyrArgPheHisIle-HisThrProLeuValLeuThrGlyLysMet----- 85
QY 507 CAAAGCCATCATACGCTCTTCCTGTATGGGAAACCTTTCTCTACCCACCGCGATGAGA 566
Db 86 -----G-----G-----G-----G-----G-----G----- 86
QY 567 AGATGGAAATTCATCTACGCTCTCCACACACTGGTGGTCAATTTCCAGAAATGGGACGA 626
Db 86 uLysGlyLeuSerSerThr-----IleArgValValGlyHis 98
QY 627 TGTTCAGTGAAGTTCTGTGAACACAGCAATGACACTTGGCGCTCAACTCATGGAA 686
Db 98 sValProGlyGluPhe-----ProVal-SerValTrpValThr-----AlaA 112
QY 687 GTGACTGTCTPACAGACAGACATGGACGGGCGCATATGTTCCCATCGCACAAAGTGAAGTGTG 746
Db 112 laAspCys-----Trp-----MetCys-----GlnProValA 121
QY 747 TAGTGGT-----AACAGATCAGATTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 794
Db 121 laArgGlyPheValValLeuProIleThrGluPheLeuValGlyAspLeuValValThrG 141
QY 795 AACGA-----TCGAATTCATCCGACGAAACCTTCCCAAGATCTCC 836
Db 141 InAsnThrSerLeuProTrpProSerSerTyrLeuThrLysThrVal-----LeuL 158
QY 837 CCATTATGTTGATGCTCTGATTCATGATCCTAGCCACTTCTCAATTTATTTACCATTA 896
Db 896 -----G-----G-----G-----G-----G-----G-----

```

```

Db 158 ysValSerPhe-----LeuLeuHisAspProSerAsnPheLeuLysThrAlaLeuPheL 176
QY 897 ACTACAAATGGAGCTTCGGGATAATACTAGCGCTCTGTTTTCACCAATCATACTGTGA 956
Db 176 euTyrSerTrpAspPheGlyAspGlyThrGlnMet---ValThrGluAspSerValValT 195
QY 957 ATCACAGTATGTCTCAATGGAACTTCAGCCTTAACCTCACTGTGTAAGACTGCAGCAC 1016
Db 195 yTrTyAsnTyrSerIleIleGlyThrPheThrValLysLeuLysValValAla----- 212
QY 1017 CAGGACCTTGTCCGCCACCGCCACCACCACCGACCTTCAAAACCCACCCCTTCTTTAG 1076
Db 212 ----- 212
QY 1077 GACCTGCTGTGACACCCCTCGAGCTGAGTAGATT---CCTGATGAAACCTGCCAGA 1133
Db 213 -----GluTrpGluGluValGluProAspAlaThrArgAlaV 225
QY 1134 TT---AACAGATATGGCCACTTTCAAGCCACCATCAAAATGTAGAGGAATCTTAGAGG 1190
Db 225 allLysGlnLysThrGlyAspPheSerAlaSerLeuLysLeuGlnGluThrLeuArgGlyI 245
QY 1191 TTAACATCATCCAGATGACAGACGTCCTGATGCCGGTGCCTGCGCTGAAAGCTCCCTAA 1250
Db 245 leGlnValLeuGlyProThr-----LeuIleGlnThrPheGlnLysMetThrValThrL 263
QY 1251 TAGACTTTTGTGACCTGCCAAGGAGCATTTCCACGAGGCTCTGTACCATCATCTTCTGT 1310
Db 263 euAsnPheLeu-----GlySerProProLeuThrValCysTrpArgLeuLys- 278
QY 1311 ACCCCACTCGGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGATGAGATGT 1370
Db 279 --ProGluCysLeuProLeuGluGluGlyGlyCysHisProValSerValAlaSerThrA 298
QY 1371 GTCTGTGACTGTGAGACGAACTTCAATGGGTCTGGGACGCTACTGTGTGAACCTCACCC 1430
Db 298 la---TyrAsnLeuThrHisThrPheArgAspProGlyAspTyrCysPheSerIleArgA 317
QY 1431 TGGGGGATGACACAAGCTGGCTCTCACGAGCACCCCTGATTTCTGTT---CCTGCACAGAG 1487
Db 317 laGluAsnIleIleSerLysThrHisGlnTyrHisLysIleGlnValTrpProSerArg- 336
QY 1488 ACCCAGCCTCGCCTTTAAGATGGCAACAGTGCCTGATCTCCGTGTGGTGTGGTGGCCA 1547
Db 337 -----IleGlnProAlaValPheAlaPheProCysAlaThrL 349
QY 1548 TATTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACACAAAGGAATACACCCCAATAG 1607
Db 349 euIleThrValMetLeuAlaPheIleMetTyr-MetThrLeuArgAsnAlaThrGlnGln 368
QY 1608 AAA 1610
Db 369 Lys 369

```

Search completed: September 10, 2004, 15:48:29  
Job time : 243 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Inc.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: September 10, 2004, 15:32:29 ; Search time 41.5 Seconds  
(without alignments)  
6620.564 Million cell updates/sec

Title: US-10-039-272-1  
Perfect score: 4850  
Sequence: 1 cggcacgagggccagagga.....aaaaaaaaaaaaaaaaaaaaa 2661

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 778828

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+, n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_p/US10039272/runat\_07092004\_144224\_21675/app\_query.fasta\_1.2823  
-DB=Issued Patents AA -QMT=fastcan -SUFFIX=ra1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTEXT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10039272.cgn 1 48 @runat\_07092004\_144224\_21675 -NCPU=6 -ICPU=3  
-NO WMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2999	61.8	560	1	US-09-985-799-90
2	2999	61.8	560	1	US-08-594-031-90
3	2999	61.8	560	4	US-09-643-597-225
4	2999	61.8	560	4	US-09-480-884A-225
5	2999	61.8	560	4	US-09-542-615A-225
6	2999	61.8	560	4	US-09-606-421B-225
7	2170	44.7	574	3	US-09-383-586-36
8	2152	44.4	572	4	US-09-197-970B-5
9	950.5	19.6	376	1	US-09-985-799-100
10	950.5	19.6	376	1	US-09-985-799-102
11	950.5	19.6	376	1	US-08-594-031-100
12	950.5	19.6	376	1	US-08-594-031-102

13	593	12.2	661	2	US-08-417-174-121	Sequence 121, App
14	593	12.2	661	3	US-09-267-439-121	Sequence 121, App
15	593	12.2	661	4	US-08-388-852B-2	Sequence 2, Appli
16	593	12.2	661	4	US-09-073-138-121	Sequence 121, App
17	592	12.2	661	2	US-08-417-174-27	Sequence 27, Appl
18	592	12.2	661	2	US-08-231-565A-27	Sequence 27, Appl
19	592	12.2	661	2	US-09-007-961-27	Sequence 27, Appl
20	592	12.2	661	3	US-09-267-439-27	Sequence 27, Appl
21	592	12.2	661	4	US-09-073-138-27	Sequence 27, Appl
22	574.5	11.8	668	1	US-07-891-942G-6	Sequence 6, Appli
23	238.5	4.9	190	4	US-08-388-852B-35	Sequence 35, Appl
24	236	4.9	202	4	US-08-388-852B-38	Sequence 38, Appl
25	233.5	4.8	192	4	US-08-388-852B-37	Sequence 37, Appl
26	233	4.8	197	4	US-08-388-852B-36	Sequence 36, Appl
27	124	2.6	71	1	US-09-985-799-98	Sequence 98, Appl
28	124	2.6	71	1	US-08-594-031-98	Sequence 98, Appl
29	111	2.3	525	4	US-09-976-594-64	Sequence 64, Appl
30	107.5	2.2	429	4	US-09-252-991A-31611	Sequence 31611, A
31	107	2.2	459	4	US-09-538-414-2	Sequence 2, Appli
32	107	2.2	459	4	US-10-074-279-2	Sequence 2, Appli
33	107	2.2	610	1	US-07-821-717B-6	Sequence 6, Appli
34	107	2.2	610	1	US-08-119-262B-6	Sequence 6, Appli
35	107	2.2	610	1	US-08-135-929A-11	Sequence 11, Appl
36	107	2.2	610	1	US-08-234-265A-11	Sequence 11, Appl
37	106.5	2.2	427	3	US-08-705-771-16	Sequence 16, Appl
38	106.5	2.2	427	4	US-09-417-540-16	Sequence 16, Appl
39	106.5	2.2	742	4	US-09-489-039A-7998	Sequence 7998, Ap
40	106	2.2	862	1	US-08-325-267A-4	Sequence 4, Appli
41	106	2.2	894	3	US-08-362-525-22	Sequence 22, Appl
42	106	2.2	894	3	US-08-971-692-15	Sequence 15, Appl
43	106	2.2	1537	1	US-08-325-267A-2	Sequence 2, Appli
44	104	2.1	865	4	US-09-281-766-19	Sequence 19, Appl
45	104	2.1	4302	3	US-08-658-136-5	Sequence 5, Appli

## ALIGNMENTS

### RESULT 1

US-09-985-799-90  
; Sequence 90, Application US/09985799  
; Patent No. RE38392  
; GENERAL INFORMATION:  
; APPLICANT: THOMPSON, Timothy C.  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES  
; NUMBER OF SEQUENCES: 175  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BAKER & BOTTS, L.L.P.  
; STREET: 1299 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004-2400

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/985,799  
FILING DATE: 06-NOV-1996  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/594,031  
FILING DATE: 30-JAN-1996

APPLICATION NUMBER: 60/006,838  
FILING DATE: 16-NOV-1995

ATTORNEY/AGENT INFORMATION:  
NAME: Remenick, James  
REGISTRATION NUMBER: 36,902

REFERENCE/DOCKET NUMBER: 0A146-0110  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-639-7700  
TELEFAX: 202-639-7890

TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 90:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 560 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; SEQUENCE DESCRIPTION: SEQ ID NO: 90:  
US-09-985-799-90

Alignment Scores:  
Pred. No.: 1,866-300 Length: 560  
Score: 2999.00 Matches: 559  
Percent Similarity: 99.82% Conservative: 0  
Best Local Similarity: 99.82% Mismatches: 1  
Query Match: 61.84% Indels: 1  
DB: 1 Gaps: 0

US-10-039-272-1 (1-2661) x US-09-985-799-90 (1-560)

QY	60	ATGGAATGCTCTACTATTCTCGGATTTCTGCTCGCTGCTGCAAGATTGCGCACTTGAT	119
DB	1	MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp	20
QY	120	GCGGCAACGATTTTCATGATGCTGGGCAATGAAGACCTTCTGCTTACATGAGGGAG	179
DB	21	AlaAlaLysArgPheHisaspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu	40
QY	180	CACAAATCAATTAATGGCTGCTTCTGATGAAATGATCGAAAGAAATCTTACCCA	239
DB	41	HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro	60
QY	240	GHTGGAGCGGGGACATAGGTGGAAACTCTGGAAGGAGCGCTGTCAGGCG	299
DB	61	ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla	80
QY	300	GTCCTGACCATGACTCACGAGCCCTCGTGGGCTCAATATATAACATTTGCGGTGACCTG	359
DB	81	ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu	100
QY	360	ATATTCCTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGACTGC	419
DB	101	IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys	120
QY	420	AGAAATGAGCTGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGCTCAGAG	479
DB	121	ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu	140
QY	480	GACAGTACGGGGAAATGGCCCGCCCAAGCCATCAATACCTCTTCCCTGTATGGGAAA	539
DB	141	AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys	160
QY	540	CTTTTCTCCACACCCCGGATGGAGAAGTGAATTTATCTACGTCTTCCACACACTT	599
DB	161	ProPheProHisProGlyTrpArgTrpAsnPheIleTyrValPheHisThrLeu	180
QY	600	GGTCAGTATTTCCAGAAATGGGACCATGTTTCACTGAGAGTTTCTGTGAACACAGCCAAT	659
DB	181	GlyClnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn	200
QY	660	GTGACACTGGGCTCAACTCATCGAAGTCACTGTCTACAGAAGACATGACGGGCATAT	719
DB	201	ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgHisGlyArgAlaTyr	220
QY	720	GTTCCCATCGCAGAAAGTGAAGATGCTAGCTGGTAAACAGATCAGATTCTCTGTTTGTG	779
DB	221	ValProIleAlaGlnValLysAspValTyrValThrAspGlnIleProValPheVal	240

RESULT 2  
US-08-594-031-90  
; Sequence 90, Application US/08594031  
; Patent No. 5783182  
; GENERAL INFORMATION:  
; APPLICANT: THOMPSON, Timothy C.  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES  
; NUMBER OF SEQUENCES: 175  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: BAKER & BOTTS, L.L.P.  
 STREET: 1299 Pennsylvania Avenue, N.W.  
 CITY: Washington  
 STATE: DC  
 COUNTRY: USA  
 ZIP: 20004-2400

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq Version 1.5  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/594,031  
 FILING DATE: 30-JAN-1996  
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/006,838  
 FILING DATE: 16-NOV-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Remenick, James  
 REGISTRATION NUMBER: 36,902  
 REFERENCE/DOCKET NUMBER: 08146-0110  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-639-7700  
 TELEFAX: 202-639-7890  
 TELEX:

INFORMATION FOR SEQ ID NO: 90:

SEQUENCE CHARACTERISTICS:

LENGTH: 560 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

ANTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

ORIGINAL SOURCE:

US-08-594-031-90

Alignment Scores:

Pred. No.: 1,86e-300 Length: 560  
 Score: 2999.00 Matches: 559  
 Percent Similarity: 99.82% Conservative: 0  
 Best Local Similarity: 99.82% Mismatches: 1  
 Query Match: 61.84% Indels: 1  
 DB: 1 Gaps: 0

US-10-039-272-1 (1-2661) x US-08-594-031-90 (1-560)

QY 60 ATGAATGCTCTACTATTCTCTGGGATTTCTGCTGGCTGCAAGATGCCACTTGAT 119  
 Db 1 MetGluCysLeuTyrPheLeuGlyPheLeuLeuAlaAArgLeuProLeuAAsp 20  
 QY 120 GCGCCCAACCAATTCATGATGTCTGGCAATGAAAGACCTTCTGCTTACATCAGGAG 179  
 Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40  
 QY 180 CACATCAATTAATGCTGTCTTCTGATGAAATGACTGGAATGAAACTCTACCCA 239  
 Db 41 HisAsnGlnLeuAsnGlyTyrPserSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60  
 QY 240 GTGTGGAGCGGGAGACATCAGGTGGGAAACTCTCTGGAGAGGAGCCGTGTGACGCG 299  
 Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla 80  
 QY 300 GTCTGTACCACTGACTCACAGCCCTCGTGGGCTCAAAATATAACATTTGGGGTGAACCTG 359  
 Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnLeuThrPheAlaValAsnLeu 100  
 QY 360 ATATTCCTTGTAGTCCAAAGAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419  
 Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnLeuValTyrGluLysAsnCys 120

QY 420 AGAATGAGGCTGGTTTATCTGCTCATCCATATGTTTACAACTGACAGCATGTCAGAG 479  
 Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpTrpAlaTrpSerGlu 140  
 QY 480 GACAGTACGGGGGAAATGGCACCGCCAAAGCCATCAACGCTCTTCCCTGATGGGAAA 539  
 Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160  
 QY 540 CTTTTCTCTCACACCCCGGATGGAGAAGATGGAAATTTTCATCTAGTCTTCCACACACTT 599  
 Db 161 PropheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180  
 QY 600 GGTGATTTTCCAGAAATTTGGAGATGTTTCAGTGAGAGTTTCTGTGAACACACCCAA 659  
 Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200  
 QY 660 GTGACACTTGGGCTCAACTCATGGAGTACTCTTACAGAACATGAGGGGATAT 719  
 Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgHisGlyArgAlaTyr 220  
 QY 720 GTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAACAGATCAGATTCTCTGTGTTGTG 779  
 Db 221 ValProIleAlaGlnValLysAspValTyrValThrAspGlnIleProValPheVal 240  
 QY 780 ACTATGTTCCAGAAAGACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838  
 Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260  
 QY 839 ATTATGTTGATGCTTCGATTCATGATCCTAGCCACTTCTCAATTTATTCACATTAAC 898  
 Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280  
 QY 899 TACAAGTGGAGCTTCGGGGATAATACCTGGGCTGTTTCTTCCACCAATCATCTGTGAAT 958  
 Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300  
 QY 959 CACAGTATGCTCAATGGAACCTTCAGCCTTAACCTCTGTGAAAGCTGCAGCACCA 1018  
 Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320  
 QY 1019 GGACCTTGTCCGCCACCGCCACACACCCAGACTTCAAAACCCACCCCTCTTTAGGA 1078  
 Db 321 GlyProCysProProProProProProProProProProProProProProProPro 340  
 QY 1079 CCTGCTGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGTGAGAACTGCCAGATTAAC 1138  
 Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360  
 QY 1139 AGATATGGCACTTCAAGCCACCATCAATTTGAGAGGGAATCTTAGAGGTTAAACATC 1198  
 Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380  
 QY 1199 ATCCAGATGACAGACGCTCTGATGCGCGTGGCTGGCTGAAAGCTCCCTAATAGACTTT 1258  
 Db 381 IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe 400  
 QY 1259 GTGCTGACCTGCCAAGGGAGCAATTCACAGGAGTCTGTACCATCATTTCTGACCCACCC 1318  
 Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420  
 QY 1319 TGGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGGATGAGATGTGTCTCTGTG 1378  
 Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440  
 QY 1379 ACTGTGAGAGACCTTCAATGGGCTGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1438  
 Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460  
 QY 1439 GACACAGGCTGTGCTCTCAGAGACCCCTGATTTCTGTTCTCTCAGACAGAGACCCAGCTCG 1498  
 Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480  
 QY 1499 CTTTTAAGGATGGCAACAGTGCCTGTATCTCGCTTGGCTGTGCTTGGCCATATTGTCACT 1558

```
Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
Qy 1559 GTGATCTCCCTCTGGTGTACAAAACACAAAGAAATACACCAATAGAAAATAGTCTCT 1618
Db 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520
Qy 1619 GGGATGTGCTCAGAACAAAGCGCTGAGTCTCTTCTCAACCGTGCAAAAGCGGTGTC 1678
Db 521 GlyAsnValAlaArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
Qy 1679 TTCGGGAAACCCAGAAAGGATCCGCTACTCTCAAAAACCAAGAAATTTAAAGGAGTTTCT 1738
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

RESULT 3
US-09-643-597-225
; Sequence 225, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Panger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455CII
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-643-597-225

Alignment Scores:
Pred. No.: 1,86e-300 Length: 560
Score: 2999.00 Matches: 559
Percent Similarity: 99.82% Conservative: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 61.84% Indels: 1
DB: 4 Gaps: 0

US-10-039-272-1 (1-2661) x US-09-643-597-225 (1-560)
Qy 60 ATGGAATGCTCTACTATTCTCTGGGATTTCTCTCTGCTGCAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuLeuAlaAlaArgLeuProLeuAsp 20
Qy 120 GCCGCCAAACGATTTCATGATGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
Qy 180 CACATCAATTAATGGCTGCTTCTGATGAATGACTGGAAATGAAACCTCTACCCA 239
Db 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGlnLysLeuTyrPro 60
Qy 240 GTCTGGAAGCGGGACATGAGTGGAAAAAATCTCTGGAAGGGAGCCGCTGTGCGAGCG 299
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla 80
Qy 300 GTCTGTACCAAGTACTCACAGCCCTCGTGGCTCAAAATATACATTTGGGGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
```

```
Qy 360 ATATTCCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTTATGAGAAGACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
Qy 420 AGAAATGAGGCTGCTTTATCTGTGCTGATCCATATGTTTACAACATGGACAGCATGCTCAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
Qy 480 GACAGTGACGGGGAAAATGGCACCGCCAAAGCCATCATACGCTCTTCCCTGATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
Qy 540 CCTTTTCTCCACCCCGGATGGAGAGATGGAATTCATCTACGTCTTCACACACTT 599
Db 161 ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180
Qy 600 GGTGAGTATTTCCAGAAATTTGGGAGGATGTTTCAGTGAGAGTTCTCTGTAACACAGCCAA 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
Qy 660 GTGACACTTGGGCTCAACTCATGGAAGTGAAGTGTCTTACAGAAGACATGGAGGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
Qy 720 GTTCCCATCGCACAAAGTGAAGATGTGACGTGGTAAACAGATCAGATTCCTGTGTTGTG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
Qy 780 ACTATGTTCCAGAAAGATCGAAATTCATCCGAGCAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
Qy 839 ATTATGTTTGTGCTCTGATTCATCATCTACCTACTTCTCAATATTCTACCATTAAC 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
Qy 899 TACAAAGTGAGCTTCGGGGATAATACTGCTGTTGTTGTTTCCCAAAATCATCTGTGAAT 958
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
Qy 959 CACAGTATGTCTCAATGGAACCTTCAGCCTTAACCTCAGTGTGAAAGCTCCAGACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
Qy 1019 GGACCTTGTCCGCCACCGCCACACACACAGACCTTCAAAAACCCACCTCTTTTGA 1078
Db 321 GlyProCysProProProProProProProProProProSerLysProThrProSerLeuGly 340
Qy 1079 CCTGCTGTGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGTGATGAAACATGCGAGATTA 1138
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
Qy 1139 AGATATGGCCACTTCAAGCCACCATCACAATGTAGAGGATCTTAGAGTTACATC 1198
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
Qy 1199 ATCCAGATCACAGACCTCTGATGCGGTGCCATGGCTGAAAGCTCCCTAATAGACTTT 1258
Db 381 IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe 400
Qy 1259 GTGCTGACCTGCCAAGGGAGCATTCACCGAGGTCTGTACCATCATTTCTTGACCCACC 1318
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420
Qy 1319 TGGAGATCACCCAGAACACAGTCTCGACCCCTGTGGATGTGATGATGATGCTGTCTGCTG 1378
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
Qy 1379 ACTGTGAGACGAACCTTCAATGGGTCTGGGACCTACTGTGTGAACCTCACCTCGGGGAT 1438
Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460
```

QY 1439 GACACAGCTGCTCTCACGAGCACCTGATTTCTGTTCTCTGACGAGACCCAGCTCG 1498  
Db |||||  
461 AspThrSerLeuAlaLeuThrSerThrLeuLeSerValProAspArgAspProAlaSer 480  
QY 1499 CTTTAAAGATGCAACAGTGCCTGATCTCGTTGGCTGCTGGCCATATTGTCACT 1558  
Db |||||  
481 ProLeuArgMetAlaAsnSerAlaLeuLeSerValGlyCysLeuAlaIlePheValThr 500  
QY 1559 GTGATCTCCCTCTGTTGTACAAAAACACAGGAATACACCAATAGAAAATAGTCCT 1618  
Db |||||  
501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520  
QY 1619 GGAATGCTGTCAGAACCAAGCTGCTGCTCTCAACCGTCGCAAAAGCCGCTTC 1678  
Db |||||  
521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540  
QY 1679 TTCCCGGGAACACAGAAAGATCGCTACTCAAAACCAAAATTAAGAGTTTCT 1738  
Db |||||  
541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560  
RESULT 4  
US-09-480-884A-225  
; Sequence 225, Application US/09480884A  
; Patent No. 6482597  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Hosken, Nancy A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C6  
; CURRENT APPLICATION NUMBER: US/09/480,884A  
; CURRENT FILING DATE: 2001-08-27  
; NUMBER OF SEQ ID NOS: 330  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 225  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-480-884A-225

Alignment Scores:  
Pred. No.: 1,86e-300 Length: 560  
Score: 2999.00 Matches: 559  
Percent Similarity: 99.82% Conservative: 0  
Best Local Similarity: 99.82% Mismatches: 1  
Query Match: 61.84% Indels: 1  
DB: 4 Gaps: 0

US-10-039-272-1 (1-2661) x US-09-480-884A-225 (1-560)

QY 60 ATGAATGCTCTACTATTCTCGGGATTCTGCTCTCGTGCAAGATTGCCACTTGAT 119  
Db |||||  
1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuLeuAlaAlaArgLeuProLeuAsp 20  
QY 120 GCCGCCAAAGCATTTTCATGATGCTCGGCAATGAAGACCTTCTGCTTACATGAGGAG 179  
Db |||||  
21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40  
QY 180 CACAATCAATTAATGCTGCTCTCTGATGAAATGACATGGATGAAACTTACCCA 239  
Db |||||  
41 HisAsnGlnLeuAsnGlyTyrSerSerAspGluAsnAspTyrAsnGluLysLeuTyrPro 60  
QY 240 GTGTGGAGCGGGAGACATGAGGTGGAATACTCTCGAAGGAGGCCGCTGTGCGAGCG 299  
Db |||||  
61 ValTyrLysArgGlyAspMetArgTyrLysAsnSerTyrLysGlyGlyArgValGlnAla 80  
QY 300 GTCTTCACCATGACTCACCAGCCCTCGTGGCTCAAAATAACATTTGCGGTGAACCTG 359  
Db |||||  
81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100

QY 360 ATATTCCCTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419  
Db |||||  
101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120  
QY 420 AGAATAGAGCTGGTTTATCTGCTGATCCATATGTTTCAACATGGACAGCATGTCAGAG 479  
Db |||||  
121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTyrThrAlaTyrSerGlu 140  
QY 480 GACAGTACGGGAAATGGCACCGCCAAAGCCATCAACGTCCTCCCTGATGGAAA 539  
Db |||||  
141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160  
QY 540 CTTTTCCTCACCCCGGATGGAGAGATGGAATTTTCATCTACGCTCTTCCACACACTT 599  
Db |||||  
161 ProPheProHisHisProGlyTyrPargArgTyrPasnPheIleTyrValPheIleThrLeu 180  
QY 600 GTCAGTATTTCAGAAATTTGGAGCATGTTTCAGTGAGAGTTTCTGTGAACACAGCAAT 659  
Db |||||  
181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200  
QY 660 GTGACACTGGGCCCTCAACTCATGAGTACTGTCTACAGAGATGAGACATGAGCGGATAT 719  
Db |||||  
201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220  
QY 720 GTTCCATGCGCACAAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCCTGTGTTTGG 779  
Db |||||  
221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240  
QY 780 ACTATGTTCCAGAGAACGATCGAAATTCATCCGACGAAAGCTTCCC-AAAGATCTCCC 838  
Db |||||  
241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260  
QY 839 ATTATGTTGATGCTCTGATTCATGATCATGATCCTAGCCATTCCTCAATATTCTACCAAT 898  
Db |||||  
261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280  
QY 899 TACAAGTGGAGCTTCGGGGATAATCTGGCTGTTTGTTCACCAATCATCTGTGAT 958  
Db |||||  
281 TyrLysTyrSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300  
QY 959 CACACGATGTGCTCAATGGAACTTCAGCTTAACTCACTGTGAAAGCTGCAGCACCA 1018  
Db |||||  
301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320  
QY 1019 GGACCTTGTCCGCCACCGCACACACACACACACCTTCAAAACCCACCTCTTTAGGA 1078  
Db |||||  
321 GlyProCysProProProProProProProProProProProProProProProProPro 340  
QY 1079 CTTGCTGTGACAAACCCCTCGGAGCTGAGTAGGATTCCTGATGAAACTGCAGATTAAAC 1138  
Db |||||  
341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360  
QY 1139 AGATATGGCCACTTTCAGCCACCATCACAATTTGTAGAGGAATCTTAGAGTTAAATC 1198  
Db |||||  
361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380  
QY 1199 ATCCAGATGACAGACTCTCTGATGCGGGTCCCATGCTGCTGAAAGCTCCCTAATAGACTTT 1258  
Db |||||  
381 IleGlnMetThrAspValLeuMetProValProTyrProTyrProTyrProTyrProTyrPro 400  
QY 1259 GTCGTACCTGCCAAGGGAGCATTTCCACGGAGGTCTGACCATCATTTCTGACCCACC 1318  
Db |||||  
401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420  
QY 1319 TCGGAGATCACCCAGAACACAGTCTCGACCCCTGTGGATGTGGATGAGATGTGTCTGCTG 1378  
Db |||||  
421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440  
QY 1379 ACTGTGAGACGACCTTCAATGGTCTGGAGCGTACTGTGTGAACCTCACCTCGGGGAT 1438  
Db |||||  
441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460  
QY 1439 GACACAAGCTTGGCTCTCACGAGCACCTGATTTCTGTTCTCTGACAGACCCAGCTCG 1498

461 AspThrSerLeuAlaLeuThrSerThrLeuLeuSerValProAspArgAspProAlaSer 480  
 1499 CCTTTAAGGATGGCAACAGTCCCTGATCTCCGTGGCTGCTTGGCCATATTTGTCACT 1558  
 481 ProLeuArgMetAlaAsnSerAlaLeuLeuSerValGlyCysLeuAlaIlePheValThr 500  
 1559 GTGATCTCCCTTGGTGTACAAACACAGAGGATACACCCCAATAGAAATAGTCTCT 1618  
 501 ValIleSerLeuLeuValTyrIleGlyHisIleGlyValProIleGluAsnSerPro 520  
 1619 GGGATGCTGGTCCAGAGCAAGCCCTGAGTGTCTTCTCAACCGTGCAAAAGCCGTGTC 1678  
 521 GlyAsnValValArgSerIleGlyLeuSerValPheLeuAsnArgAlaIleValPhe 540  
 1679 TTCCCGGGAACACGAGAAAGGATCCGCTACTCAAAACCAAGAAATTTAAGAGTCTCT 1738  
 541 PheProGlyAsnGlnGluIleAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

RESULT 5

US-09-542-615A-225  
 ; Sequence 225, Application US/09542615A  
 ; Patent No. 6518256

GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong  
 ; APPLICANT: Fan, Liqun  
 ; APPLICANT: Kalos, Michael D.  
 ; APPLICANT: Bangur, Chaitanya S.  
 ; APPLICANT: Hosken, Nancy A.  
 ; APPLICANT: Fanger, Gary R.  
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
 ; FILE REFERENCE: 210121.455C8  
 ; CURRENT APPLICATION NUMBER: US/09/542, 615A  
 ; CURRENT FILING DATE: 2000-04-14  
 ; NUMBER OF SEQ ID NOS: 350  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 225  
 ; LENGTH: 560  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapien  
 US-09-542-615A-225

Alignment Scores:  
 Pred. No.: 1.86e-300 Length: 560  
 Score: 2999.00 Matches: 559  
 Percent Similarity: 99.82% Conservative: 0  
 Best Local Similarity: 99.82% Mismatches: 1  
 Query Match: 61.84% Indels: 1  
 DB: 4 Gaps: 0

US-10-039-272-1 (1-2661) x US-09-542-615A-225 (1-560)

60 ATGGAAGTCTCTACTATTCTCTGGGATTTCTGCTCGTGGTCAAGATTGCCACTTGAT 119  
 1 MetGluCysLeuTyrIlePheLeuGlyPheLeuLeuLeuAlaAlaArgLeuProLeuAsp 20  
 120 GCGCCCAACAGATTTCATGATGTCTGGCAATGAAAGACCTTCTGCTTACATGAGGAG 179  
 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40  
 180 CACATCAATTAATGCTGGTCTCTCTGATGAAATGCTGGAATGAAATCTTACCCTA 239  
 41 HisAsnGlnLeuAsnGlyIleTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60  
 240 GTGTGGAGCGGGAGACATGAGTGGGAAACCTCTGGAAGGAGGCGGTGTGACGCG 299  
 61 ValIleLysArgGlyAspMetArgTrpLysAsnSerIlePheGlyArgValGlnAla 80  
 300 GTCTTGACCAAGTACTACACCGCTCTGCGGCTCAATATATAACATTTGCGGTGACCTG 359  
 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100

360 ATATTCCTAGATCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGACTGC 419  
 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120  
 420 AGAATAGAGCTGCTTTATCTGCTGATCCCATATGTTTCAACTGGACAGCATGGTCAGAG 479  
 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140  
 480 GACAGTGACGGGGAAAAATGGCACCGGGCCAAAGCCATATAACGCTTCCCTGATGGGAAA 539  
 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160  
 540 CCTTTTCTCCACCCCGGATGGAGAAGATGGAAATTTCTACTACGTCTTCCACACACTT 599  
 161 ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180  
 600 GGTGAGTATTTCCAGAAATTTGGAGCATGTTTCAGTGAGAGTTCCTGTGAAACACAGCCA 659  
 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200  
 660 GTGACACTTTGGGCCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGGACGGCATAT 719  
 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220  
 720 GTTCCCATCGCAAGTGAAGATGTGTACGTGTGTAACAGATCAGATTCCTCTGTTTGTG 779  
 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240  
 780 ACTATGTTCCAGAGAAGCATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838  
 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260  
 839 ATTATGTTTGATGTCCTGATTCATGATCCTACGCCATCTCCTCAATATTCTACCATTAAC 898  
 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280  
 899 TCAAGTGAGCTTGGGGATAATACTGGCTGTTTGTTCACCAACATCATACTGTGAAT 958  
 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300  
 959 CACACGTATGTCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGACCA 1018  
 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320  
 1019 GGAACCTTGTCCGCCACCGCCACCAACCCAGACCTTCAAAACCCACCCCTCTTTAGGA 1078  
 321 GlyProCysProProProProProProProProProProProProProProProProPro 340  
 1079 CTTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAATCCAGATTAAC 1138  
 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnGlnIleAsn 360  
 1139 AGATATGGCCACTTTCAAGCCACCATCACAAATGTAGAGGAATCTTAGAGGTTAACATC 1198  
 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380  
 1199 ATCCAGATGACAGACGTCCTGATGCCGGTCCCATGGCTGAAAGCTCCCTAATAGACTTT 1258  
 381 IleGlnMetThrAspValLeuMetProValProProProProProProProProProPro 400  
 1259 GTCTGTGACCTGCCAAGGGAGCATTCACCGAGGTCTGTACCATCATTTCTGACCCACACC 1318  
 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420  
 1319 TGCAGATACCCAGAACACAGTCTGCAGCGCTCTGGATGTGGATGATGATGTGTGTGTG 1378  
 421 CysGlnIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440  
 1379 ACTGTGAGACCAACCTTCAATGGGTCTGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1438  
 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460  
 1439 GACACAGCCTGGCTCTCACAGACACCTCTGATTCTTCTGTTCTCTGACAGAGACCCGCTCG 1498





Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460  
Qy 1439 GACACAGCCTGGCTCTCACAGACACCTGATTTCTGTTCTGACAGAGACCCAGCTCG 1498  
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480  
Qy 1499 CCTTTAAGGATGGCAACAGTCCCTGATCTCCGTTGCTGCTGGCCATATTTGTCACT 1558  
Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500  
Qy 1559 GTGATCTCCCTCTTGGTGTCAAAAACACAGGAATACACCCCAATAGAAATAGTCTCT 1618  
Db 501 VallieSerLeuLeuValTyrLysIleShisIysGluTyrAsnProIleGluAsnSerPro 520  
Qy 1619 GGAATGTGTCAGAGCAAGGCTGATGCTCTTCTCAACCGTGCAAGCCGTGTC 1678  
Db 521 GlyAsnValValArgSerIleGlyLeuSerValPheLeuAsnArgAlaIleValPhe 540  
Qy 1679 TTCCCGGAAACACAGAAAAGGATCCGCTACTCAAAAACCAAGAAATTTAAAGGAGTTTCT 1738  
Db 541 PheProGlyAsnGlnGluIleAspProLeuLeuIleYsAsnGlnGluPheIleGlyValSer 560

RESULT 7  
US-09-383-586-36  
; Sequence 36, Application US/09383586  
; Patent No. 6242419  
; GENERAL INFORMATION:  
; APPLICANT: Strachan, Lorna  
; APPLICANT: Sleeman, Matthew  
; APPLICANT: Abernethy, Nevin  
; APPLICANT: Onrust, Rene  
; APPLICANT: Kumble, Anand  
; APPLICANT: Murison, Greg  
; TITLE OF INVENTION: Compounds isolated from stromal cells  
; TITLE OF INVENTION: and methods for their use  
; FILE REFERENCE: 11000.1037c1  
; CURRENT APPLICATION NUMBER: US/09/383,586  
; CURRENT FILING DATE: 1999-08-26  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 36  
; LENGTH: 574  
; TYPE: PRT  
; ORGANISM: Mouse  
US-09-383-586-36

Alignment Scores:  
Pred. No.: 6,87e-215 Length: 574  
Score: 2170.00 Matches: 402  
Percent Similarity: 82.63% Conservative: 69  
Best Local Similarity: 70.53% Mismatches: 83  
Query Match: 44.74% Indels: 17  
DB: 3 Gaps: 2

US-10-039-272-1 (1-2661) x US-09-383-586-36 (1-574)

Qy 60 ATGGAATGCTCTACTATTCTCGGATTTCTGCTGCTGCTGCAAGATTGCCACTTGAT 119  
Db 1 MetGluSerLeuCysGlyValLeuGlyPheLeuLeuLeuAlaIleGlyLeuProLeuGln 20  
Qy 120 GCCGCCAAACGATTTTCATGATGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179  
Db 21 AlaAlaIleArgPheArgAspValLeuGlyHisGluGlnTyrProAsnHisMetArgGlu 40  
Qy 180 CACAATCAATAATGGCTGCTCTTCTGATGAAATGACTGGAATGAAACCTCTACCCA 239  
Db 41 HisAsnGlnLeuArgGlyTyrSerSerAspGluAsnGluTyrAspGluHisLeuTyrPro 60  
Qy 240 GTGTGGAAGCGGGAGACATGAGTGGAAAACTCTCTGGAAGGAGGCGCTGTGCGAGCG 299  
Db 61 ValTrpArgArgGlyAspGlyArgTrpIleSerAspSerTrpGluGlyArgValGlnAla 80  
Qy 300 GTCTGTGACCAGTACTCACCAGCCCTCGTGGGCTCAAAATATAACATTTGCGGTGAACCTG 359

Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100  
Qy 360 ATATTTCCTAGATGCCAAAGGAAGATGCCAATGCAACATAGTCTATGAGAAGAACTGC 419  
Db 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120  
Qy 420 AGAAATGAGGCTGGTTTATCTGTCATCCATATGTTTACAACCTGGACAGCATGGTCAGAG 479  
Db 121 ArgAsnAspLeuGlyLeuThrSerAspLeuHisValTyrAsnTrpThrAlaGlyAlaAsp 140  
Qy 480 GACAGTGACGGGGAAATATGGACCGCCAAAGCCATCATACGCTTCCCTCATCGGAAA 539  
Db 141 AspGlyAspTrpGluAspGlyThrSerArgSerGlnHisLeuArgPheProAspArgArg 160  
Qy 540 CCTTTTCCTCACCACCCCGGATGGAGAAGATGGAATTTCACTCTCTTCACACACTT 599  
Db 161 PropheProArgProHisGlyTyrIleLysIleSerPheValTyrValPheHisThrLeu 180  
Qy 600 GGTCAGTATTTCCAGAAATGGGACGATGTTTCAGTGAGAGTTTCTGTGACACACGCCAAT 659  
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerAlaArgValSerIleAsnThrValAsn 200  
Qy 660 GTGACACTTGGCCCTCAACTCATGGAAGTACTGTCTACAGAGACATCGAGCGGCATAT 719  
Db 201 LeuThrAlaGlyProGlnValMetGluValThrValPheArgArgTyrGlyArgAlaTyr 220  
Qy 720 GTTCCCATCGCACAAAGTGAAGATGTCAGTGGTAAACAGATCAGATTCCTCTGTTTGG 779  
Db 221 IleProIleSerLysValIleAspValTyrValIleThrAspGlnIleProValPheVal 240  
Qy 780 ACTATGTTCCAGAGAAGACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCC 838  
Db 241 ThrMetSerGlnLysAsnAspArgAsnLeuSerAspGluIlePheLeuArgAspLeuPro 260  
Qy 839 ATTATGTTTGATGCTCTGATTCATCATCTAGCCACTCTCTCAATTTATCTACCATTAAC 898  
Db 261 IleValPheAspValIleHisAspProSerHisPheLeuAsnAspSerAlaIleSer 280  
Qy 899 TACAAAGTGAGCTTCGGGATAATCTGGCCCTGTTGTTTCCACCAATCATCTGTGAAT 958  
Db 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300  
Qy 959 CACAGTATGCTCTCAATGGAACCTTCAGCCCTTAACCTCCTGCTGGAACCTCGAGACCA 1018  
Db 301 HisThrTyrValLeuAsnGlyThrPheAsnLeuAsnLeuThrValGlnThrAlaValPro 320  
Qy 1019 GGACCTGTGTCG-----CCACGCGCACCA 1042  
Db 321 GlyProCysProProProSerProSerThrProProProSerThrProProSerPro 340  
Qy 1043 CCACCCAGACCT-----TCAAAACCCACCCCTTCTTTAGGACCTGCTGCTGAC 1090  
Db 341 ProProSerProLeuProThrLeuSerThrProSerProSerLeuMetProThrGlyTyr 360  
Qy 1091 AACCCCTCGAGCTGAGTAGGATTCCTGATGAAACTGCCAGATTAAACAGATTATGCCAC 1150  
Db 361 LysSerMetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyr 380  
Qy 1151 TTTCAGCCACCATCATCAATCTGAGAGGAATCTTAGAGTTAAACATCATCCAGATGACA 1210  
Db 381 PheArgAlaThrIleThrIleValGluGlyIleLeuGluValSerIleMetGlnIleAla 400  
Qy 1211 GACCTCTCATGTCGCGTGCATGGCCCTGAAAGCTCCCTAATAGACTTTTGTGCTGACTGC 1270  
Db 401 AspValProMetProThrProGlnProAlaAsnSerLeuMetAspPheThrValThrCys 420  
Qy 1271 CAAGGAGCATTCCTCCCGAGGCTGTGATCCATCTATTTCTGACCCACCTCGGAGATCAC 1330  
Db 421 LysGlyAlaThrProMetGluAlaCysThrIleIleSerAspProThrCysGlnIleAla 440  
Qy 1331 CAGAACACAGTCTGACGCCCTGGGATGTGATGAGATGTCTGCTGACTGTGAGACGA 1390



Db 441 GlnAsnArgValCysSerProValAlaValAspGlyLeuCysLeuLeuSerValArgArg 460  
 QY 1391 ACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCTGGGGGATCACACAGCCTG 1450  
 Db 461 AlaPheAsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAlaSerLeu 480  
 QY 1451 GCTCTCAGAGACACCTGATTTCTGTCTGTGACAGAGACCCAGCCTCGCTTTAAGGATG 1510  
 Db 481 AlaLeuThrSerThrLeuLeuSerIleProGlyLysAspProAspSerProLeuArgAla 500  
 QY 1511 GCAACAGTCCCTGATCTCGTGGCTGCTGGCCATATTTGCTCACTGTGATCTCCCTC 1570  
 Db 501 ValAsnGlyValLeuLeuSerIleGlyCysLeuAlaValLeuValThrMetValThrIle 520  
 QY 1571 TTGGTGTCAAAAACACAAAGAAATACAAACCAATAGAAAATAGTCTGGGAATGTGTC 1630  
 Db 521 LeuLeuTyrLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 540  
 QY 1631 AGAAGCAAGGCTGAGTGTCTTCTCAACCGTGCAGAAAGCCGTGTCTTCCCGGAAAC 1690  
 Db 541 LysGlyLysGlyLeuSerValLeuLeuSerHisAlaLysAlaProPhePheArgGlyAsp 560  
 QY 1691 CAGGAAAGGATCCGCTACTCAAAACCAA 1720  
 Db 561 GlnGluLysAspProLeuLeuGlnAspLys 570

## RESULT 8

US-09-197-970B-5  
 ; Sequence 5, Application US/09197970B  
 ; Patent No. 6664385  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Michele Sanicola-Nadel  
 ; Joseph V. Bonventre  
 ; Catherine A. Hession  
 ; Takaharu Ighimura  
 ; Henry Wei  
 ; Richard L. Cate  
 ;  
 ; TITLE OF INVENTION: MODULATORS OF TISSUE REGENERATION  
 ; NUMBER OF SEQUENCES: 7  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Biogen, Inc.  
 ; STREET: 14 Cambridge Center  
 ; CITY: Cambridge  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02142  
 ;  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/197,970B  
 ; FILING DATE: 23-No. 6664385-1998  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 60/018,228  
 ; FILING DATE: 24-MAY-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Levine, Leslie M.  
 ; REGISTRATION NUMBER: 35,245  
 ; REFERENCE/DOCKET NUMBER: A010 PCT CIP  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 679-2810  
 ; TELEFAX: (617) 679-2838  
 ; INFORMATION FOR SEQ ID NO: 5:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 572 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Alignment Scores:  
 Pred. No.: 4,94e-213 Length: 572  
 Score: 2152.00 Matches: 395  
 Percent Similarity: 82.5% Conservative: 74  
 Best Local Similarity: 69.54% Mismatches: 85  
 Query Match: 44.37% Indels: 15  
 DB: 4 Gaps: 1

US-10-039-272-1 (1-2661) x US-09-197-970B-5 (1-572)

QY 60 ATGGAAATGCTCTACTATTTCTGGGATTTCTGCTCTGCTGCTGCAAGATTGCCACTGTAT 119  
 Db 1 MetGluSerLeuCysGlyValLeuValPheLeuLeuLeuAlaAlaGlyLeuProLeuGln 20  
 QY 120 GCGGCCAAACGATTTTCATGATGTGCTGGCGCAATGAAGACCTTCTGCTTACATGAGGAG 179  
 Db 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40  
 QY 180 CACATCAATTAATGGCTGCTCTTCTGATGAAATGACTGGAATGAAAACTTACCCA 239  
 Db 41 AsnAsnGlnLeuArgGlyTrpSerSerAspGluAsnGluTrpAspGluGlnLeuTyrPro 60  
 QY 240 GTGTGGAAGCGGGAGACATGAGGTGGAATACTCTCGAAGGAGGCGCTGTGCAGCG 299  
 Db 61 ValTrpArgArgGlyGluGlyArgTrpLysAspSerTrpGluGlyGlyArgValGlnAla 80  
 QY 300 GTCCTGACCACTGACTCACAGCCCTCGTGGCTCAATATATAACATTTCCGCTGACCTG 359  
 Db 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100  
 QY 360 ATATCCCTAGATGCCAAAGAAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419  
 Db 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120  
 QY 420 AGAATGAGGCTGTTTATCTGCTGATCCATATGTTTACAACTGCAGCAGATGTCAGAG 479  
 Db 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTrpThrThrGlyAlaAsp 140  
 QY 480 GACAGTCAGCGGGAATGGCCGCGCAAGCCATCATACGCTTCCCTGATGGGAAA 539  
 Db 141 AspGluAspTrpGluAspSerThrSerGlnGlyHisLeuArgPheProAspGlyLys 160  
 QY 540 CTTTCTCACCACCCCGGATGGAGAAGATGGAATTTCTACGTCTTCCCTGATGGGAAA 599  
 Db 161 ProPheProArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeu 180  
 QY 600 GGTCACTATTCCAGAAATTTGGGACGATGTTCACTGAGAGTTTCTGTGAACACAGCCAA 659  
 Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerAlaArgValSerIleAsnThrValAsn 200  
 QY 660 GTGACACTTGGGCTCAACTCATGGAGTCACTGTCTACAGAGACATGACCGGCAATAT 719  
 Db 201 LeuThrValGlyProGlnValMetGluValIleValPheArgHisGlyArgGalaTyr 220  
 QY 720 GTTCCCATCCCAAGTGAAGATGTGTACGTGGTAACAGATCAGATTCCTGTGTTTGTG 779  
 Db 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal 240  
 QY 780 ACTATGTTCCAGAGAACGATCGAAATTCATCCGAGAAACCTTCCC-AAAGATCTCCCC 838  
 Db 241 ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260  
 QY 839 ATTATGTTTGTGCTCTGATTCATGATTCCTAGCCACTTCTCAATATTCTACCAATAC 898  
 Db 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280  
 QY 899 TACAAGTGAGCTTCGGGATAATACTGGCTGTTTGTGTTTCCACCAATCATACTGTGAAT 958  
 Db 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300  
 QY 959 CACACGTATGTCTCAATGGAACCTTCAGCTTAACCTCACTGTGAAAGCTGCAGACCA 1018

301 HisThrTyrValLeuLeuAsnGlyThrPheAsnLeuThrValGlnThrAlaValPro 320  
1019 GGACTTTGTCGCCACCGCCACACCCAGACT----- 1054  
321 GlyProCysProSerProThrProSerProSerSerSerSerProAlaSer 340  
1055 -----TCAAAACCCACCCCTCTTTAGACCTGCTGGTGACAACCCC 1096  
341 SerProSerProThrLeuSerThrProSerProSerLeuMetProThrGlyHisLysSer 360  
1097 CTGAGCTGAGTAGGATTCCTGATGAAACCTGCCAGATTAACAGATATGCCACTTCA 1156  
361 MetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyThrPheArg 380  
1157 GCCACCATCAATTTAGAGGAATCTTAGAGTTAATACATCATTCAGATGACAGACGTC 1216  
381 AlaThrIleThrIleValAspGlyIleLeuGluValAsnIleGlnValAlaAspVal 400  
1217 CTGATGCGGTCGCCATGCTGAAAGCTCCCTAATAGACTTTGTCTGACCTGCCAAGG 1276  
401 ProIleProThrProGlnProAsnSerLeuMetAspPheIleValThrCysLysGly 420  
1277 AGCATTCACCGAGGTCTGTACCATCTTCTGACCCACCTGCGAGATCACCCAGAAC 1336  
421 AlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAlaGlnAsn 440  
1337 ACAGTCTGAGCCCTGTGATGTGATGATGTGTCTGCTGCTGCTGAGACGACCTTC 1396  
441 ArgValCysSerProValAlaValAspGluLeuCysLeuLeuSerValArgAlaPhe 460  
1397 ATGGGTCTGGACGCTGCTGTGACCTCACCTGGGGGATGACACAGCCCTGCTCTC 1456  
461 AsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAlaSerLeuAlaLeu 480  
1457 ACCAGCACCTGATTTCTGCTCTGACAGACCCAGCCCTCGCTTAAAGATGGCAAC 1516  
481 ThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerProLeuArgThrValAsn 500  
1517 AGTGCCTGATCTCCCTGGCTGCTGGCCATATTTGCTGCTGCTGCTGCTGCTGCTG 1576  
501 GlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMetValThrIleLeuLeu 520  
1577 TACAAAACACCAAGCAATACACCAATAGAAATAGTCTGGGAATGTGTGAGAAGC 1636  
521 TyrLysLysHisLysThrTyrLysProIleGlyAsnCysThrArgAsnValValLysGly 540  
1637 AAAGGCTGAGTGTCTTTCTCAACCGTGCACCAAGCCGTGTTCTCCCGGGAACACGAA 1696  
541 LysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgGlyAspArgGlu 560  
1697 AGGATCCGCTACTCAAAACCAA 1720  
561 LysAspProLeuLeuGlnAspLys 568

RESULT 9

US-09-985-799-100  
; Sequence 100, Application US/09985799  
; Patent No. RE38392  
; GENERAL INFORMATION:  
; APPLICANT: THOMPSON, Timothy C.  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES  
; NUMBER OF SEQUENCES: 175  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BAKER & BOTTS, L.L.P.  
; STREET: 1299 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004-2400  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS

SOFTWARE: FastSEQ Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/985,799  
FILING DATE: 06-NOV-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/594,031  
FILING DATE: 30-JAN-1996  
APPLICATION NUMBER: 60/006,838  
FILING DATE: 16-NOV-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Remenick, James  
REGISTRATION NUMBER: 36,902  
REFERENCE/DOCKET NUMBER: 0A146-0110  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-639-7700  
TELEFAX: 202-639-7890  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 100:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 376 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
SEQUENCE DESCRIPTION: SEQ ID NO: 100:  
US-09-985-799-100  
Alignment Scores:  
Pred. No.: 4e-89 Length: 376  
Score: 950.50 Matches: 299  
Percent Similarity: 59.31% Conservative: 26  
Best Local Similarity: 54.56% Mismatches: 46  
Query Match: 19.60% Indels: 177  
DB: 61 Gaps: 61  
US-10-039-272-1 (1-2661) x US-09-985-799-100 (1-376)  
QY 99 GTGCAAGATGCCACTTGCATGCGCCAAACGATTTCTGCTGGCATGAAGA 158  
Db 6 AlaAlaArg-----AspAlaAlaLysArg---HisAspVal---GlyAsnArg--- 19  
QY 159 CCTTCTGCTTACATGAGGAGCAGCACATCAATTAATATGCTGCTTCTCATGAAAAATGAC 218  
Db 20 ---SerAlaLysMetArg---HisAsn-----AsnGlyTrpSerSerAsp---AsnAsp 34  
QY 219 TGGATGAAAAAATCTTACCCAGTGTGGAAGCGGGAGACATGAGGTGGAAAAAATCTCTGG 278  
Db 35 TrpAsnLysTyr-----ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrp 51  
QY 279 AAGGAGCGCGTGTGAGCGGCTCTGACGAGTACTCACCACCTCTCGTGGGCTCAAT 338  
Db 52 LysGlyGlyArgVal-----AlaValThrSerAspSer-----AlaValGlySerAsn 67  
QY 339 ATAACATTTTCGGTGAACCTGTATTCCTTAGATGCCAAAGGAAGATGCCAATGGCAAC 398  
Db 68 Thr-----AlaValAsn-----ArgCys-----LysAspAlaAsnGlyAsn 79  
QY 399 ATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGTTTCTCTGCTCATCATGTTTAC 458  
Db 80 ValTyr-----LysAsnCysArgAsn---AlaGly---SerAlaAsp---TyrValTyr 94  
QY 459 AACTGGACAGCATGCTGAGGAGACAGTACGCGGGAATAATGGACCGCCCAAGCATCAT 518  
Db 95 AsnTrpThrAlaTrpSer---AspSerAspGly---AsnGlyThrGly---SerHisHis 111  
QY 519 AAGCTCTTCTTCATGGAACCTTTTCTCCTCACCACCCCGGATGAGAGATGGAATTC 578  
Db 112 AsnVal-----AspGlyLys-----HisHisGlyTrpArgTrpAsnTyr 125

QY 579 ATCTACGCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGATGTTCACTGAGA 638  
Db : : : : :  
126 ValHisThrGlyTyrLys-----GlyArgCysSerValArg 137  
QY 639 GTTCTGTGAACACACGCAATGTGACACTTGGGCGCTCAACTCATGGAAGTGAAGTCTCTAC 698  
Db : : : : :  
138 ValSerValAsnThrAlaAsnValThr-----GlyMetValThrValTyr 152  
QY 699 AGACACATGACGGGCAATGTTCCCATCGCACAAAGTGAAGATGTGTAGTGTGAACA 758  
Db : : : : :  
153 ArgArgHisGlyArgAlaTyrVal-----AlaValLysAspValTyrValValThr 169  
QY 759 GATCAGATTCCTGTGTTGTGCTACTATGTTCCAGAGACGATCGAAATTCATCCGACGAA 819  
Db : : : : :  
170 Asp-----ValValThrMet-----LysAsnAspArgAsnSerSerAsp--T 183  
QY 819 ACCTTCCCAAGATCTCCCATTTATGTTGATGTCCTGATTCATGATCTAGCCACTTCC 878  
Db : : : : :  
183 hr-----LysAspMetAsp-----ValHisAsp---SerHis----- 192  
QY 879 TCATATTCTACCAATTAACATACAGTGGAGCTTCGGGGATAAATACTGGCCTGTTGTTT 938  
Db : : : : :  
193 --AsnTyrSerThr---AsnTyrLysTrpSer---GlyAspAsnThrGly-----ValS 208  
QY 939 CCACCAATCATACTGTGAATCACACGATGCTCAATGGAACCTTCAGCCCTTAACCTCA 998  
Db : : : : :  
208 erThrAsnHisThrValAsnHisThrTyrVal---AsnGlyThrSerAsn-----T 224  
QY 999 CTGTGAAGCTGCAGCACCAGAGACTTGTCCGCCACCGCCACCACCCAGACCTTCAA 1058  
Db : : : : :  
224 hrValLysAlaAlaAla-----GlyCys----- 231  
QY 1059 AACCCACCCCTTCTTAGGACCTGCTGGTGCACACCCCTGGAGTGAAGTAGGATTCCTG 1118  
Db : : : : :  
232 ----ArgSerLysThrSerGlyAlaGlyAspAsn-----SerArg----- 243  
QY 1119 ATGAAACTGCACATTAACAGATATGCCACTTTCAGCCACCACCATCACATTTAGAGG 1178  
Db : : : : :  
244 --AspAsnCys-----AsnArgTyrGlyHis-----AlaThrThrValGly- 256  
QY 1179 GAATCTTAGAGTTAATCATCATCATGATGACAGACGTCCTGATCCCGTGCCATGGCCTG 1238  
Db : : : : :  
257 ----ValAsnMetThrAspValMetVal-----TrpSerS 267  
QY 1239 AAGCTCCCTAATAGACTTGTGTGACCTGCCAAGGAGGATTTCCACGAGGTCGTGA 1298  
Db : : : : :  
267 erAsp-----ValValThrCys---GlySer-----T 275  
QY 1299 CCATCAATTCGACCCACCTGGAGATCACCAGACACAGTCTGCAGCCCTGTGGATG 1358  
Db : : : : :  
275 hrValCysThrSerAspThrCys-----ThrAsnThrValCysSer---ValAspV 291  
QY 1359 TGATGATGATGTGTCTGCTGACTGTGAGACGAACTTCAATGGTCTGGGACGTAAGTG 1418  
Db : : : : :  
291 aLAsp---MetCys-----ThrValArgargThr---AsnGlySerGlyThrTyrCysV 307  
QY 1419 TGAACCTCACCTGGGGATGACACAGCTGGCTCTCACAGACACCCCTGATTTCTGTT 1478  
Db : : : : :  
307 aLAsn-----ThrGlyAspAspThrSerAlaThrSerThrSerVal----- 320  
QY 1479 CTGACAGACCCAGCTCGCTTTAAGGATGCGCAACAGTGCCTGATCTCGTGGCT 1538  
Db : : : : :  
321 --AspArgAspAlaSer-----ArgMetAlaAsnSerAla-----SerValGlyC 335  
QY 1539 GCTTGGCCATTTTGTGCTGATCTCCCTCTGTGTGTACAAAAACACAGGAATACA 1598  
Db : : : : :  
335 ys-----AlaValThrValSerValTyrLysLysHisLys----- 346  
QY 1599 ACCCAATAGAAATAGTCTGGGAATGTGTGACAGCAAGCCCTGAGTGTCTTCTCA 1658  
Db : : : : :  
347 -----TyrAsnAsnSerGlyAsnValValArgSerLysGly-----SerValA 361

QY 1659 ACCGTGCAAAAGCCGTGTTCTTCCCGGGAACACGAAAGAGTCCGCTACTCAAAACC 1718  
Db : : : : :  
361 snArgAlaLysAlaVal-----GlyAsnLysAspLysAsn----- 372  
QY 1719 AAGAATTTAAAGGAGTTTCT 1738  
Db : : : : :  
373 -----LysGlyValSer 376  
RESULT 10  
US-09-985-799-102  
; Sequence 102, Application US/09985799  
; Patent No. RE38392  
; GENERAL INFORMATION:  
; APPLICANT: THOMPSON, Timothy C.  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES  
; NUMBER OF SEQUENCES: 175  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BAKER & BOTTS, L.L.P.  
; STREET: 1299 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004-2400  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/985,799  
; FILING DATE: 06-Nov. RE38392-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/594,031  
; FILING DATE: 30-JAN-1996  
; APPLICATION NUMBER: 60/006,838  
; FILING DATE: 16-NOV-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Remenick, James  
; REGISTRATION NUMBER: 36,902  
; REFERENCE/DOCKET NUMBER: 08146-0110  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-639-7700  
; TELEFAX: 202-639-7890  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 102:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 376 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; SEQUENCE DESCRIPTION: SEQ ID NO: 102:  
US-09-985-799-102  
Alignment Scores:  
Pred. No.: 4e-89 Length: 376  
Score: 950.50 Matches: 299  
Percent Similarity: 59.31% Conservative: 26  
Best Local Similarity: 54.56% Mismatches: 46  
Query Match: 19.60% Indels: 177  
DB: 1 Gaps: 61  
US-10-039-272-1 (1-2661) x US-09-985-799-102 (1-376)  
QY 99 GCTGCAAGATTGCCACTTGCATCGCCCAACGATTTTCATGATGCTGCTGGCATGAAGA 158  
Db 6 AlaAlaArg-----AspAlaAlaLysArg---HisAspVal---GlyAsnArg--- 19

```
QY 159 CTTCTCTGTTACATGAGGACACAAATCAATTAATGCTGCTCTTCTGATGAAATGAC 218
Db 20 ---SerAlaThrMetArg---HisAsn-----AsnGlyThrSerSerAsp---AsnAsp 34
QY 219 TGGAAATGAAAACTCTACCCAGTGTGGAAGGGGAGACATGAGTGGAAAAAATCTCTGG 278
Db 35 TrpAsnLysTyr-----ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrp 51
QY 279 AAGGAGGCGGTGTGACGGCGTCTGACCAAGTACTCACCAGCCCTGCTGGGTCAAAAT 338
Db 52 LysGlyGlyArgVal-----AlaValThrSerAspSer-----AlaValGlySerAsn 67
QY 339 ATAACATTTGGGTGAACCTGATATTCCTAGATGCCAAAGGAGATGCCAATGGCAAC 398
Db 68 Thr-----AlaValAsn-----ArgCys-----LysAspAlaAsnGlyAsn 79
QY 399 ATAGTCTATGAGAAGAACTCAGAGAAATGAGGCTGGTTTATCTGTGATCCATATGTTAC 458
Db 80 ValTyr-----LysAsnCysArgAsn---AlaGly---SerAlaAsp---TyrValTyr 94
QY 459 AACTGGACAGCATGTCAGAGACAGTACGCGGGAATGACCGCGCAAGCCATCAT 518
Db 95 AsnTrpThrAlaTrpSer---AspSerAspGly---AsnGlyThrGly---SerHisHis 111
QY 519 AAGGCTTCCCTGATGGGAAACCTTTTCTCACCACCCCGGATGGAGAGATGGAATTC 578
Db 112 AsnVal-----AspGlyLys-----HisHisGlyTrpArgTrpAsnTyr 125
QY 579 ATCTACGCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGACGATGTTTCAGTGAGA 638
Db 126 ValHisThrGlyTyrLys-----GlyArgCysSerValArg 137
QY 639 GTTCTGTGAAACAGCCAAATGTGACACTTGGGCGCTCACTCATGGAAGTGACTCTAC 698
Db 138 ValSerValAsnThrAlaAsnValThr-----GlyMetValThrValTyr 152
QY 699 AGAAGACATGGAGCGGCATATGTTCCATCGCACAGTGAAGAGATGTGTACGTGTAAACA 758
Db 153 ArgArgHisGlyArgAlaTyrVal-----AlaValLysAspValTyrValValThr 169
QY 759 GATCAGATCTCTGTGTTGTGACTATGTTCCAGAGAACGATCGAAATTCACCGACGAA 818
Db 170 Asp-----ValValThrMet-----LysAsnAspArgAsnSerSerAsp--T 183
QY 819 ACCTTCCCAAGATCTCCCAATATGTTGTGATCTCTGATTCATGATCTCAGCCATCC 878
Db 183 hr-----LysAspMetAsp-----ValHisAsp---SerHis----- 192
QY 879 TCAATTTATCTACCAATTAACACAGTGGAGCTTCGGGATAATACTGGCTGTGTTGTTT 938
Db 193 ---AsnTyrSerThr---AsnTyrLysTrpSer---GlyAspAsnThrGly-----ValS 208
QY 939 CCACCAATCATACTGTGAATCATACAGTATGTGTCAATGGAACCTTCAGCCTTAACCTCA 998
Db 208 erThrAsnHisThrValAsnHisThrTyrVal---AsnGlyThrSerAsn-----T 224
QY 999 CTGTGAAGCTGCAGACACAGGACCTGTGCGGCCACCGCCACCACCCAGACCTTCAA 1058
Db 224 hrValLysAlaAlaAla-----GlyCys----- 231
QY 1059 AACCCACCCCTCTTTAGGACCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTG 1118
Db 232 -----ArgSerLysThrSerGlyAlaGlyAspAsn-----SerArg----- 243
QY 1119 ATGAAACATGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATCATCAATTTAGAGG 1178
Db 244 --AspAsnCys-----AsnArgTyrGlyHis-----AlaThrThrValGly- 256
QY 1179 GAATCTTAGAGGTTAATCATCATCCAGTACAGACGCTCTGTGATGCCGGTCCCATGGCCTG 1238
Db 257 -----ValAsnMetThrAspValMetVal-----TrpSerS 267
QY 1239 AAAGTCCCTAATAGACTTTGTGCTGACCTGCGCAAGGGAGCATTTCCACGGAGTCTGTA 1298
```

```
Db 267 erAsp-----ValValThrCys-----GlySer-----T 275
QY 1299 CCATCATTTCTGACCCACCTGCGAGATCACCCAGAAACACAGTCTTGACGCGCTGTGGATG 1358
Db 275 hrValCysThrSerAspThrCys-----ThrAsnThrValCysSer---ValAspV 291
QY 1359 TGGATGAGATGTGTCTGCTGACTGTGAGACGAAACCTTCAATGGGTCTGGGACGTACTGTG 1418
Db 291 alAsp---MetCys-----ThrValArgArgThr---AsnGlySerGlyThrTyrCysV 307
QY 1419 TGAACCTCACCTCGGGGATGACACAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTC 1478
Db 307 alAsn-----ThrGlyAspAspThrSerAlaThrSerThrSerVal----- 320
QY 1479 CTGACAGAGACCCAGCCTCGCTTTAAGGATGGCAACACAGTGCCTGTATCTCCGTTGGCT 1538
Db 321 --AspArgAspAlaSer-----ArgMetAlaAsnSerAla-----SerValGlyC 335
QY 1539 GCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACACAAAGGAATACA 1598
Db 335 ys-----AlaValThrValSerValTyrLysLysHisLys----- 346
QY 1599 ACCCAATGAAAAATAGTCTCGGAATGTGTCAGAGCAAAAGGCTGAGTGTCTTCTCA 1658
Db 347 -----TyrAsnAsnSerGlyAsnValValArgSerLysGly-----SerValA 361
QY 1659 ACCGTGCAAAAGCCTGTTCTCCGCGGAAACACAGGAAAGGATCCGCTACTCAAAAACC 1718
Db 361 snArgAlaLysAlaVal-----GlyAsnLysAspLysAsn----- 372
QY 1719 AAGAAATTTAAAGGAGTCTTCT 1738
Db 373 -----LysGlyValSer 376

RESULT 11
US-08-594-031-100
; Sequence 100, Application US/08594031
; Patent No. 5783182
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Timothy C. IDENTIFYING METASTATIC SEQUENCES
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; FILING DATE: 30-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/006,838
; FILING DATE: 16-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Remenick, James
; REGISTRATION NUMBER: 36,902
; REFERENCE/DOCKET NUMBER: 0A146-0110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-639-7700
; TELEFAX: 202-639-7890
; TELEX:
; INFORMATION FOR SEQ ID NO: 100:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
```



```

; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/594,031
; FILING DATE: 30-JAN-1996
; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/006,838
; FILING DATE: 16-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Remenick, James
; REGISTRATION NUMBER: 36,902
; REFERENCE/DOCKET NUMBER: 0A146-0110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-639-7700
; TELEFAX: 202-639-7890
; TELEX:
;
; INFORMATION FOR SEQ ID NO: 102:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; US-08-594-031-102

```

```

Alignment Scores:
Pred. No.: 4e-89
Score: 950.50
Percent Similarity: 59.31%
Best Local Similarity: 54.56%
Query Match: 19.60%
DB: 1

```

US-10-039-272-1 (1-2661) x US-08-594-031-102 (1-376)

```

QY 99 GCTGCAAGATGTCACCTGATGCGCGCAACGATTTTCATGATGCTGGCCATGAAGA 158
Db 6 AlaAlaArg-----AspAlaAlaLysArg---HisAspVal---GlyAsnArg--- 19
QY 159 CCTTCTGCTTACATGAGGGAGCCAAATCAATTAAATGGCTGGTCTTCTGATGAATGAC 218
Db 20 ---SerAlaTyrMetArg---HisAsn-----AsnGlyTyrSerSerAsp---AsnAsp 34
QY 219 TGAATGAAATACTCTACCCAGTGTGGAAGCGGGGAGACATGAGGTGGAAAACTCTCTGG 278
Db 35 TrpAsnLysTyr-----ValTyrLysArgGlyAspMetArgTyrLysAsnSerTyr 51
QY 279 AAGGAGCGCGTGTGAGCGGCTCTGACAGAGTACTCACCAGCCCTCTGGGGCTCAAT 338
Db 52 LysGlyGlyArgVal-----AlaValThrSerAspSer-----AlaValGlySerAsn 67
QY 339 ATACATTTGGGTGAACCTGATATTCCTAGATGCCAAAGGAAGATGCCAATGGCAAC 398
Db 68 Thr-----AlaValAsn-----ArgCys-----LysAspAlaAsnGlyAsn 79
QY 399 ATAGCTATGAGAACTGCGAGAAATGAGGCTGGTTATCTGTGTATGATGATGTTTAC 458
Db 80 ValTyr-----LysAsnCysArgAsn---AlaGly---SerAlaAsp---TyrValTyr 94
QY 459 AACTGGACAGATGTCAGAGACAGTGCACGGGAAATGCGACCGCCCAAGCCATCAT 518
Db 95 AsnTrpThrAlaTyrSer---AspSerAspGly---AsnGlyThrGly---SerHisHis 111
QY 519 AACGTCCTTCCTGATGGGAAACCTTTTCTCACCACCCCGGATGGAGATGGAATTC 578

```

```

Db 112 AsnVal-----AspGlyLys-----HisGlyTyrArgTyrAsnTyr 125
QY 579 ATCTACGCTTCCACACACTGGTTCAGTATTTCAGAAATGGACAGATGTTTCAGTGAGA 638
Db 126 ValHisThrGlyTyrLys-----GlyArgCysSerValArg 137
QY 639 GTTTCGTGAACACAGCAATGTGACACTTGGGCTCAACTCATCTGGAAGTACTGTCTAC 698
Db 138 ValSerValAsnThrAlaAsnValThr-----GlyMetValThrValTyr 152
QY 699 AGAAGACATGACGCGGATATGTTCCCATCCACAGTGAAGATGTCTAGCTGTATAACA 758
Db 153 ArgArgHisGlyArgAlaTyrVal-----AlaValLysAspValTyrValValThr 169
QY 759 GATCAGATTCTGTGTTTGTGACTATGTTCCAGAAAGACGATCGAAATTCATCCGACGAA 818
Db 170 Asp-----ValValThrMet-----LysAsnAspArgAsnSerSerAsp--T 183
QY 819 ACCTTCCCAAGATCTCCCATTTATGTTTGTGATGCTCTGATTCATCATCTACCCACTCC 878
Db 183 hr-----LysAspMetAsp-----ValHisAsp---SerHis----- 192
QY 879 TCAATTATTCTACCATTAACACTACAGTGGAGCTTCGGGATATAACTGCGCTGTTGTTT 938
Db 193 --AsnTyrSerThr--AsnTyrLysTyrSer--GlyAspAsnThrGly-----ValS 208
QY 939 CCACCAATCATCTGTGAATCACAGTATGCTCAATGGAACCTTTCAGCCCTTAACCTCA 998
Db 208 erThrAsnHisThrValAsnHisThrTyrVal--AsnGlyThrSerAsn-----T 224
QY 999 CTGTGAAAGCTGCAGCACGACGACCTTGTCCGCCACCGCCACCCACCCAGACCTTCA 1058
Db 224 hrValLysAlaAlaAla-----GlyCys----- 231
QY 1059 AACCCACCCCTCTTTAGGACCTGTGTTGTCGACAAACCCCTGGAGCTGAGTAGGATTCCTG 1118
Db 232 -----ArgSerLysThrSerGlyAlaGlyAspAsn-----SerArg----- 243
QY 1119 ATGAAACCTGCGCAGATTAACAGATATGGCCACTTTCAGGCCACCATCAAAATGTAGAG 1178
Db 244 --AspAsnCys-----AsnArgTyrGlyHis-----AlaThrThrValGly- 256
QY 1179 GAATCTTAGAGGTTAATCATCATCCAGATCACAGACCTCTGTGATGCGGTGCCATGGCCTG 1238
Db 257 -----ValAsnMetThrAspValMetVal-----TrpSerS 267
QY 1239 AAGCTCCCTAATAGACTTGTGTCACCTGCCAAGGAGCATTTCCACGGAGTCTGTA 1298
Db 267 erAsp-----ValValThrCys---GlySer-----T 275
QY 1299 CCATCATTTCTGACCCACCTCGGAGATCACCCAGAACACAGTCTGCGCCCTGTGATG 1358
Db 275 hrValCysThrSerAspThrCys-----ThrAsnThrValCysSer---ValAspV 291
QY 1359 TGGATGAGATGTCCTCTGTCGACTGTGAGACACCTTCAATGGTCTGGGACGACTGTG 1418
Db 291 alaAsp--MetCys-----ThrValArgArgThr--AsnGlySerGlyThrTyrCysV 307
QY 1419 TGAACCTCACCTGGGGGATGACACAGCCTGGCTCTCACGAGCACCCCTGATTCTGTTTC 1478
Db 307 alaAsn-----ThrGlyAspAspThrSerAlaThrSerThrSerVal----- 320
QY 1479 CTGACAGACCCAGCCTCGCTTTAAGAGTGGCAACAGTGCCTGATCTCGCTTGGCT 1538
Db 321 --AspArgAspAlaSer-----ArgMetAlaAsnSerAla-----SerValGlyC 335
QY 1539 GCTTGGCCATATTGTTCACGTGTGATCTCCCTCTTGTGTACAAAACACACAGGAATACA 1598
Db 335 ys-----AlaValThrSerValTyrLysLysHisLys----- 346
QY 1599 ACCCAATAGAAAATAGTCTGGGAATGTGGTCAAGCAAGAGGCTGAGTGTCTTCTTCA 1658

```



COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/267,439  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,174  
FILING DATE: 05-APR-1995  
APPLICATION NUMBER: US/08/231,565  
FILING DATE: 22-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: CAROL M. GRUPPI  
REGISTRATION NUMBER: 37,341  
REFERENCE/DOCKET NUMBER: 2026-4124US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 751-8800  
TELEFAX: (212) 751-8849  
TELEX: 421792  
INFORMATION FOR SEQ ID NO: 121:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 661  
TYPE: amino acid  
STRANDEDNESS: Unknown  
TOPOLOGY: Unknown  
MOLECULE TYPE: Protein  
US-09-267-439-121

Alignment Scores:  
Pred. No.: 4,31e-52 Length: 661  
Score: 593.00 Matches: 167  
Percent Similarity: 39.45% Conservative: 93  
Best Local Similarity: 25.34% Mismatches: 199  
Query Match: 12.23% Indels: 201  
DB: 3 Gaps: 17

US-10-039-272-1 (1-2661) x US-09-267-439-121 (1-661)

QY 66 TGTCTCTACTATTTCTGGGATTTCTGCTCTGCTGCAAGATTGCACTTGATGCGGCC 125  
Db 8 CysLeu-----LeuHisLeuAlaValIleGlyAlaLeuLeuAlaValGlyAlaThr 24

QY 126 AAA-----CGATTTTCATGATGCTGGGCAATGAAGACCTTCTGCTACATGAGGAG 179  
Db 25 LysValProArgAsnGlnAspTrpLeuGlyValSerArg----- 37

QY 180 CACAATCAATTAATGCTGCTCTCTGATGAAATGACATGGAATGAAACCTTACCCCA 239  
Db 38 -----GlnLeuArg-----ThrLysAlaTrpAsnArgGlnLeuTyPro 50

QY 240 GTGTGGAAGCGGGGAGACATGAGTGGGAAAAAATCTCTGGAGGGAGGCGCTGTGAGGCG 299  
Db 51 GluTrpThr-----GluAlaGlnArgLeuAspCysTrpArgGlyGlyGlnValSerLeu 68

QY 300 GTCTGACACGATCACTCAGCCCTCGTGGCTCAATATATAACATTTCGGTGAACCTG 359  
Db 69 LysValSerAsnAspGlyProThrLeuIleGlyAlaAsnAlaSerPheSerIleAlaLeu 88

QY 360 ATATTCCTTAGATGCCAAAGGAAGATGCCAATGCGCAACATAGTCTATGAGAAGACTGC 419  
Db 89 AsnPheProGlySerGlnLysValLeuProAspGlyGlnValIleTrpValAsnThr 108

QY 420 AGAAATGAGGCTGTTTATCTGCTGATCCATATGTTTACAC-----TGG--- 464  
Db 109 -----IleIleAsnGlySerGlnValTrpGly 117

QY 465 -----ACAGCATGTCAGAGGACAGTACGGGGAATAATGACCCCGCCAAAGCATCAT 518  
Db 118 GlyGlnProValTyProGlnGluThrAspAspAla----- 129

QY 519 AACGCTCTCCCTGATGGAAACCTTTCTCCACCACCCCGGATGAGGAAGATGAATTC 578  
Db 130 CysIlePheProAspGlyGlyProCysProSerGlySerTrpSerGlnLysArgSerPhe 149

1039 ----- 1039  
370 GluSerThrGlyMetThrProGluLysValProValSerGluValMetGlyThrThrLeu 389  
1039 ----- 1039  
390 AlaGluMetSerThrProGluAlaThrGlyMetThrProAlaGluValSerIleValVal 409  
1039 ----- 1039  
410 LeuSerGlyThrThrAlaAlaGlnValThrThrThrGluTyrValGluThrThrAlaArg 429  
1040 -----CCACCACCAGACCTTCAAAACC----- 1063  
430 GluLeuProIleProGluProGluGlyProAspAlaSerSerIleMetSerThrGluSer 449  
1064 ---ACCCCTCTTTAGGACCTGCTGTCAGCAACCCCTGGAGCTGAGTAGGATTCCTGAT 1120  
450 IleThrGlySerLeuGlyProLeuLeuAspGlyThrAlaThrLeuArgLeuValLysArg 469  
1121 GAA-----AACTGCCAGATTAAACAGATATGCCCTTTCAAGCCACCATCAAT 1171  
470 GlnValProLeuAspCysValLeuTyArgTyrGlySerPheSerValThrLeuAspIle 489  
1172 GTAGAGGAATCTTAGAGGTTAAATCATCCAGATGACAGACGCTCTGATGCCGGTGCCA 1231  
490 ValGlnGly-----IleGluSerAlaGluIleLeuGlnAlaValPro 503  
1232 TGGCTGGAAGCTCCCTAATAGCTTTGTCGTGACCTGCCAAGGGAGCATTCCTCCAGGAG 1291  
504 SerGlyGluGlyAspAlaPheGluLeuThrValSerCysGlnGlyGlyLeuProLysGlu 523  
1292 GNTCTTACCATCATTTCTGACCCCTCGCGATCACCAGAACACATCTGAGCCCT 1351  
524 AlaCysMetGluIleSerSerProGlyCysGlnProProAlaGlnArgLeuCysGlnPro 543  
1352 GTGGATGTGGATGAGATGTCTGTGCTGATGTGAGACGACCTTC-----AATGGTCTGGG 1408  
544 ValLeuProSerProAlaCysGlnLeuValLeuValHisGlnIleLeuLysGlyGlySerGly 563  
1409 ACGTACTGTGTAACCTCACTCCCTGGGGATGACACAAAGCTGCTCTCAGACACCCCTG 1468  
564 ThrTyCysLeuAsnValSerLeuAlaAspThrAsnSerLeuAlaValValSerThrGln 583  
1469 ATTTCTGTTCTCAGACAGACCCAGCTCGCTTTAAGGATGCAACACAGTGCCTGATC 1528  
584 LeuIleMetProGlyGlnGluAlaGly-----LeuGlyGlnValProLeuIle 599  
1529 TCCGTTGGCTGTGGTGCATATTTGTCACTGTGATCTCCCTCTCTGTTGTAACAAAA 1585  
600 ValGlyIleLeuLeuValLeuMetAlaValValLeuAlaSerLeuIleTyArgArg 618

RESULT 14  
US-09-267-439-121  
Sequence 121, Application US/09267439  
Patent No. 6270778  
GENERAL INFORMATION:  
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,  
APPLICANT: STEVEN A.  
TITLE OF INVENTION: MELANOMA ANTIGENS AND  
THEIR USE IN DIAGNOSTIC AND THERAPEUTIC  
METHODS  
NUMBER OF SEQUENCES: 126  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM: FLOPPY DISK  
MEDIUM TYPE: FLOPPY DISK



Qy	66	TGTC	TACTA	TATTC	TCTGG	ATTCT	CTCTC	CTGCA	GATTG	CCAC	ATTGAT	CGCG	CC	125
Db	8	CysLeu	-----	-----	LeuHis	LeuAla	ValIle	GlyAla	LeuLeu	AlaVal	GlyAla	Thr	24	

126	AAA-----CGATTTTCATGATGCTGGCGCAATGAAGAACCTTCCTGCTTATCATGAGGAG	179
25	LysValProArgAsnGlnAspTrpLeuGlyValSerArg-----	37
180	CACAATCAATTAATGGCTGGTCTCTGATGAAATGACTGGAATGAATAAACTCTACCCA	239
38	-----GlnLeuArg-----ThrLysAlaTrpAsnArgGlnLeuTyPro	50
240	GTCTGGAAGCGGGAGACATGAGTGGMAAAATCTCTGAAGGGAGGCGGTGTGACGGCG	299
51	GlutTrpThr-----GluAlaGlnArgLeuAspCysTrpArgGlyGlnValSerLeu	68
300	GTCTGACCACTGACTCACCAGCCCTGTGGGCTCAATATATAACTTTGGCGGTGAACCTG	359
69	LysValSerAsnAspGlyProThrLeuIleGlyAlaAsnAlaSerPheSerIleAlaLeu	88
360	ATATTCTCTAGATGCCAAAAGAGATGCCAATGGCCACATAGTCTATCAGAAGAACTGC	419
89	AsnPheProGlySerGlnLysValLeuProAspGlyGlnValIleTrpValAsnAsnThr	108
420	AGAAATGAGCTGGTTTATCTGCTGCTCATATGTTTACAAC-----TGG--	464
109	-----IleIleAsnGlySerGlnValTrpGly	117
465	-----ACAGCATGCTCAGAGCAGTGCAGGGGAAAATGGCACCGGCCCAAGACCATCAT	518
118	GlyGlnProValTyProGlnIleThrAspAspAla-----	129
519	AACGCTCTTCCTGATGGGAAACCTTTTCTCCACCACCCCGGATGGAGAGATGGAATTTC	578
130	CysIlePheProAspGlyGlyProCysProSerGlySerTrpSerGlnLysArgSerPhe	149
579	ATCTACGCTTCCACACACTTGTGTCAGTATTTCAGAAAATTTGGAGCATGTTTCAGTGAGA	638
150	ValTrpValIleTrpLysThrTrpGlyGlnTyTrpGlnValLeuGlyGlyProValSerGly	169
639	GTTTCTGTGAACACAGCCCAATGTGCACACTTGGGCCTCAACTCATGGAAGTGTCTGTCTAC	698
170	LeuSerIleGlyThrGlyArgAlaMetLeuGlyThrHisThrMetGluValThrValTyTrp	189
699	AGAAGACATATGA-----CGGGCATATGTTCCCATCGCACAACTGAAAGATGTGACGTGTA	755
190	HisArgArgGlySerArgSerTyValProLeuAlaHisSerSerSerAlaPheThrIle	209
756	ACAGATCAGATTCTGTGTTTGTGACTATGTTCCAGAAAGAACGATCGAAATTCATCCGAC	815
210	ThrAspGlnValProPheSerValSerValSerGlnLeuArgAlaLeuAspGlyGlyAsn	229
816	GAAGACTTCCC-AAAGATCTCCCATATTATGTTGATGCTCTGATTCATGATCTACGCCAC	874
230	LysHisPheLeuArgAsnGlnProLeuThrPheAlaLeuGlnLeuHisAspProSerGly	249
875	TTCTCAATATTCTACATTACTACAGTGGAGCTTCGGGATAATACTGGCCTGTTT	934
250	TyrLeuAlaGluAlaAspLeuSerTyThrTrpAspPheGlyAspSerSerGlyThrLeu	269
935	GTTTCCACCAATCATCTGTCGAATCACACGATATGCTGTCATGGAACCTTCAGCCTTAAC	994
270	IleSerArgAlaLeuValValThrHisThrTyLeuGluProGlyProValThrAlaGln	289
995	CTCACTGTGAAGCTGCAGACCA-----GGACCTTGTCCGCCACCGCCA---	1039
290	ValValLeuGlnAlaAlaIleProLeuThrSerCysGlySerSerProValProGlyThr	309
1039	-----	1039
310	ThrAspGlyHisArgProThrAlaGluAlaProAsnThrThrAlaGlyGlnValProThr	329
1039	-----	1039
330	ThrGluValValGlyThrThrProGlyGlnAlaProThrAlaGluProSerGlyThrThr	349
1039	-----	1039

Search completed: September 10, 2004, 15:52:45  
Job time : 78.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2004, 01:34:28 ; Search time 213 Seconds  
(without alignments)  
6932.978 Million cell updates/sec

Title: US-10-039-272-1  
Perfect score: 2661  
Sequence: 1 cggcagggccagagga.....aaaaaaaaaaaaaaaaaa 2661

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2547.8	95.7	2669	1	US-09-985-799-91
2	2547.8	95.7	2669	1	US-09-985-799-101
3	2547.8	95.7	2669	1	US-08-594-031-91
4	2547.8	95.7	2669	1	US-08-594-031-101
5	1051.2	39.5	2303	4	US-09-197-970B-4
6	1047.8	39.4	2213	3	US-09-383-586-27
7	545.6	20.5	619	4	US-09-123-912-105
8	545.6	20.5	619	4	US-09-643-597-105
9	545.6	20.5	619	4	US-09-643-597-121
10	545.6	20.5	619	4	US-09-480-884A-105
11	545.6	20.5	619	4	US-09-480-884A-121
12	545.6	20.5	619	4	US-09-542-615A-105
13	545.6	20.5	619	4	US-09-542-615A-121
14	545.6	20.5	619	4	US-09-606-421B-105
15	545.6	20.5	619	4	US-09-606-421B-121
16	545.6	20.5	619	4	US-09-221-107-105
17	545.6	20.5	619	4	US-09-221-107-121
18	493	18.5	494	4	US-09-389-681-449
19	493	18.5	494	4	US-09-620-405B-449
20	493	18.5	494	4	US-09-433-826B-449
21	493	18.5	494	4	US-09-604-287A-449
22	493	18.5	494	4	US-09-834-759-449
23	480.8	18.1	698	3	US-09-040-384-5
24	480.8	18.1	698	4	US-09-123-912-5
25	480.8	18.1	698	4	US-09-643-597-5
26	480.8	18.1	698	4	US-09-480-884A-5
27	480.8	18.1	698	4	US-09-542-615A-5

28 480.8 18.1 698 4 US-09-606-421B-5  
29 480.8 18.1 698 4 US-09-221-107-5  
30 312.6 11.7 620 1 US-09-985-799-113  
31 312.6 11.7 620 1 US-08-594-031-113  
32 306.8 11.5 524 1 US-09-985-799-105  
33 306.8 11.5 524 1 US-08-594-031-105  
34 210.2 7.9 473 1 US-09-985-799-115  
35 210.2 7.9 473 1 US-08-594-031-115  
36 208.4 7.8 335 1 US-09-985-799-92  
37 208.4 7.8 335 1 US-09-985-799-94  
38 208.4 7.8 335 1 US-09-985-799-96  
39 208.4 7.8 335 1 US-09-985-799-123  
40 208.4 7.8 335 1 US-08-594-031-92  
41 208.4 7.8 335 1 US-08-594-031-94  
42 208.4 7.8 335 1 US-08-594-031-96  
43 208.4 7.8 335 1 US-08-594-031-123  
44 189.4 7.1 286 1 US-09-985-799-121  
45 189.4 7.1 286 1 US-08-594-031-121

## ALIGNMENTS

RESULT 1  
US-09-985-799-91  
; Sequence 91, Application US/09985799  
; Patent No. RE38392  
; GENERAL INFORMATION:  
; APPLICANT: THOMPSON, Timothy C.  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES  
; NUMBER OF SEQUENCES: 175  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BAKER & BOTTS, L.L.P.  
; STREET: 1299 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004-2400  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/985,799  
; FILING DATE: 06-NOV-1995  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/594,031  
; FILING DATE: 30-JAN-1996  
; APPLICATION NUMBER: 60/006,838  
; FILING DATE: 16-NOV-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Remenick, James  
; REGISTRATION NUMBER: 36,902  
; REFERENCE/DOCKET NUMBER: DA146-0110  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-639-7700  
; TELEFAX: 202-639-7890  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 91:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2669 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: <Unknown>  
; ORIGINAL SOURCE:  
; SEQUENCE DESCRIPTION: SEQ ID NO: 91:  
US-09-985-799-91

Query Match 95.7%; Score 2547.8; DB 1; Length 2669;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 2591; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

QY	28	AACTTGGTGGCTGGCTGGAGAAATTCAGCATGGAATGCTCTACTATTTCCTGGGAT	87	1140	GTAGATTCTTGATGAAAACCTGCCAGATTAAACAGATATGGCCACTTTTCAAGCCACCATCA	1199
Db	60	ACCTTGAGTGGCTGGCTGGAGAAATTCAGCATGGAATGCTCTACTATTTCCTGGGAT	119	1167	CAATTTGTAGAGGGAAATCTTAGAGGTAAACATCATCCAGATGACAGACGTCTCTGATGCCGG	1226
QY	88	TTCTGCTCCTGGCTGCAAGATTGGCACTTGATGCCGCCAAACGAAATTCATGATGTCTCG	147	1200	CAATTTGTAGAGGGAAATCTTAGAGGTAAACATCATCCAGATGACAGACGTCTCTGATGCCGG	1259
Db	120	TTCTGCTCCTGGCTGCAAGATTGGCACTTGATGCCGCCAAACGAAATTCATGATGTCTCG	179	1227	TGCCATGGCTTCAAAAGCTCCCTAATAGACTTTTGTGCTGACCTTGCCTGCAAGGGAGCATTTCCCA	1286
QY	148	GCAATGAAGACCTTCTGCTTTACATGAGGAGGACCAATCAATTAATGGCTGCTTCTTG	207	1260	TGCCATGGCTTCAAAAGCTCCCTAATAGACTTTTGTGCTGACCTTGCCTGCAAGGGAGCATTTCCCA	1319
Db	180	GCAATGAAGACCTTCTGCTTTACATGAGGAGGACCAATCAATTAATGGCTGCTTCTTG	239	1287	CGGAGGTCTGTACCATCATTTCTGACCCCACTGCGAGATCACCCAGAACACAGTCTGCA	1346
QY	208	ATGAAATGACTTGAATGAAATCTTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA	267	1320	CGGAGGTCTGTACCATCATTTCTGACCCCACTGCGAGATCACCCAGAACACAGTCTGCA	1379
Db	240	ATGAAATGACTTGAATGAAATCTTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA	299	1347	GCCCTGTGGATGTGGATGAGATGTGTCTGTGACTGTGAGACGAACTTTCAATGGGTCTG	1406
QY	268	AAATCTCTGGAAGGGAGCGGTGTGAGGGGTCCTGACAGTGAATCACCAGCCCTCG	327	1380	GCCCTGTGGATGTGGATGAGATGTGTCTGTGACTGTGAGACGAACTTTCAATGGGTCTG	1439
Db	300	AAATCTCTGGAAGGGAGCGGTGTGAGGGGTCCTGACAGTGAATCACCAGCCCTCG	359	1407	GGACGTACTGTGTGAACCTCACCTCGGGGGATGACACAAGCTTGGCTCTCAGAGACCC	1466
QY	328	TGGGCTCAATATACATTTTGGGCTGAACTTATTCCTAGATGCCAAAGGAGATG	387	1440	GGACGTACTGTGTGAACCTCACCTCGGGGGATGACACAAGCTTGGCTCTCAGAGACCC	1499
Db	360	TGGGCTCAATATACATTTTGGGCTGAACTTATTCCTAGATGCCAAAGGAGATG	419	1467	TGATTTCTGTTCTTGACAGAGACCCAGCTCCCTTTAAGGATGCAACACAGTGCCTCGA	1526
QY	388	CCAAATGGCAATAGTCTATGAGAGAACTGACAGAAATGAGGCTGTTTATCTGCTGATC	447	1500	TGATTTCTGTTCTTGACAGAGACCCAGCTCCCTTTAAGGATGCAACACAGTGCCTCGA	1559
Db	420	CCAAATGGCAATAGTCTATGAGAGAACTGACAGAAATGAGGCTGTTTATCTGCTGATC	479	1527	TCCTCGTGGCTTGGCCATATTTGTCACTGTGATCTCCCTTTGGTGTGTAACAAAAAC	1586
QY	448	CATATGTTTAACTGGACAGCATGGTCAAGAGACAGTCAAGGGAATGGACCGGCC	507	1560	TCCTCGTGGCTTGGCCATATTTGTCACTGTGATCTCCCTTTGGTGTGTAACAAAAAC	1619
Db	480	CATATGTTTAACTGGACAGCATGGTCAAGAGACAGTCAAGGGAATGGACCGGCC	539	1587	ACAAGGAATACAAACCCCAATAGAAAATAGTCTCGGAATGTGCTCAGAACAAAGGCTCGA	1646
QY	508	AAAGCCATCAATACGTTCTTCCGTGAGGAAACCTTTTCTCACCACCCCGGATGAGAA	567	1620	ACAAGGAATACAAACCCCAATAGAAAATAGTCTCGGAATGTGCTCAGAACAAAGGCTCGA	1679
Db	540	AAAGCCATCAATACGTTCTTCCGTGAGGAAACCTTTTCTCACCACCCCGGATGAGAA	599	1647	GTGCTCTTCTCAACCGTGCAAAAGCGTGTCTTCCCGGAAACCCAGGAAAAGGATCCCG	1706
QY	568	GATGGAATTTCACTACGCTCTTCCACACACTTGGTCACTATTTCCAGAAATGGGACGAT	627	1680	GTGCTCTTCTCAACCGTGCAAAAGCGTGTCTTCCCGGAAACCCAGGAAAAGGATCCCG	1739
Db	600	GATGGAATTTCACTACGCTCTTCCACACACTTGGTCACTATTTCCAGAAATGGGACGAT	659	1707	TACTCAAAAAACCAAGAAATTTAAAGAGTTTCTTAAATTTTCGACCTTGTCTTCAAGCTCA	1766
QY	628	GTTCACTGAGAGTTTCTGTGAACACAGCAATGTGACACTTGGGCTCAACTCATGGAAG	687	1740	TACTCAAAAAACCAAGAAATTTAAAGAGTTTCTTAAATTTTCGACCTTGTCTTCAAGCTCA	1799
Db	660	GTTCACTGAGAGTTTCTGTGAACACAGCAATGTGACACTTGGGCTCAACTCATGGAAG	719	1767	CTTTTCAGTGCCTTGAATGTGAGATGTGCTGGAGTGGCTATTAAACCTTTTTCCTAAAG	1826
QY	688	TGACTGTCTACAGAAAGACATGAGCGGCATATGTTCCCATCGCACAGTGAAGATGTGT	747	1800	CTTTTCAGTGCCTTGAATGTGAGATGTGCTGGAGTGGCTATTAAACCTTTTTCCTAAAG	1859
Db	720	TGACTGTCTACAGAAAGACATGAGCGGCATATGTTCCCATCGCACAGTGAAGATGTGT	779	1827	ATTATTTGTTAAATAGATATTTGCTTTGGGAAAGTTGAAATTTTTTATAGTTTAAATGTCA	1886
QY	748	ACGTGGTAAACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAGAAAGCATCGAAT	807	1860	ATTATTTGTTAAATAGATATTTGCTTTGGGAAAGTTGAAATTTTTTATAGTTTAAATGTCA	1919
Db	780	ACGTGGTAAACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAGAAAGCATCGAAT	839	1887	TTTTAGAGATGGGAGAGGATTAATACTGCAGGACGCTTTCAGCCATGTTTGTGAAACTGAT	1946
QY	808	CATCCGACGAAACCTTCC-CAAGATCTCCCAATATGTTTGTATGCTCTGATTCATGATC	866	1920	TTTTAGAGATGGGAGAGGATTAATACTGCAGGACGCTTTCAGCCATGTTTGTGAAACTGAT	1979
Db	840	CATCCGACGAAACCTTCC-CAAGATCTCCCAATATGTTTGTATGCTCTGATTCATGATC	899	1947	AAAAAGCACTTAGCAAGGCTCTTTTCAATATTTTTTATGTTTCACTTATAAGTCTTAG	2006
QY	867	CTAGCCACTTCTCAATATTCTTACCAATTAACCTACAGTGGAGCTTGGGGATTAATCTG	926	1980	AAAAAGCACTTAGCAAGGCTCTTTTCAATATTTTTTATGTTTCACTTATAAGTCTTAG	2039
Db	900	CTAGCCACTTCTCAATATTCTTACCAATTAACCTACAGTGGAGCTTCCGGGATTAATCTG	959	2007	GTAACTAGTAGGATAGAAACACTGTGTCCCGAGAGTAAGAGAGAGGACTACTATTGATTA	2066
QY	927	GCCTGTTTGTTCACCAATCATACTGTGAATCACAGTATGCTCAATGGAACCTTCA	986	2040	GTAACTAGTAGGATAGAAACACTGTGTCCCGAGAGTAAGAGAGAGGACTACTATTGATTA	2099
Db	960	GCCTGTTTGTTCACCAATCATACTGTGAATCACAGTATGCTCAATGGAACCTTCA	1019	2067	GAGCCTAAACCCAGGTTAACTGCAAGAGAGGCGGGATACCTTTCAGCTTTCCAGTAACTG	2126
QY	987	GCCTTAACTCTGTTGAAAGCTGACGACACAGGACCTTGTCCGCCACCCGCCACCCAC	1046	2100	GAGCCTTAAACCCAGGTTAACTGCAAGAGAGGCGGGATACCTTTCAGCTTTCCAGTAACTG	2159
Db	1020	GCCTTAACTCTGTTGAAAGCTGACGACACAGGACCTTGTCCGCCACCCGCCACCCAC	1079	2127	TATGCATAAAGCCCAATGTAGTCCAGTTTCTTAAGATCATGTTTCCAGAGCTAACTGAATCCCA	2186

Db 2160 TATGCAATAAGCCAAATAGTCCAGTTTCTAAGATCATGTTTCCAAAGTAACTGAATCCCA 2219  
QY 2187 CTTCAATACACACTCATGAACCTCTGATGAACAATAACAGGCCCAAGCCTGFGTATGA 2246  
Db 2220 CTTCAATACACACTCATGAACCTCTGATGAACAATAACAGGCCCAAGCCTGFGTATGA 2279  
QY 2247 TGTGCACACTTGTAGACTCAGAAAAAATACTACTCTATAAATGGTGGGAGTATTTG 2306  
Db 2280 TGTGCACACTTGTAGACTCAGAAAAAATACTACTCTATAAATGGTGGGAGTATTTG 2339  
QY 2307 GTGCAACACTTGTCTGGCTGAGTGAAGGAATGATATTCATATATTCATTATTC 2366  
Db 2340 GTGCAACACTTGTCTGGCTGAGTGAAGGAATGATATTCATATATTCATTATTC 2399  
QY 2367 TGGCAATTTAGTTAGTCTTTTATATACAGGCATGATCTGAGTGAACCTCTGTTG 2426  
Db 2400 TGGCAATTTAGTTAGTCTTTTATATACAGGCATGATCTGAGTGAACCTCTGTTG 2459  
QY 2427 TATTTCCAAATTTTGTATAGTCTGTCACATATTTGAATCAAAATATTAGACTTTCC 2486  
Db 2460 TATTTCCAAATTTTGTATAGTCTGTCACATATTTGAATC-ATATATTAAGACTTTCC 2518  
QY 2487 AAAAATTTGGTCTGCTTTTTCATGGCAACTTGATCAAGTAAAGGATTTCCCTCTGTTG 2546  
Db 2519 AAAGATGAGTCTGCTTTTTCATGGCAACTTGATCAAGTAAAGGATTTCCCTCTGTTG 2578  
QY 2547 GAACATAAACCAATTTACTATATGTTAGACAGACATTTTTTTTTTCTCTGTTG 2606  
Db 2579 TAACTAAACCAATTTACTATATGTTAGACATGACATTTCTTTCTCTCTGTTG 2638  
QY 2607 -AAATGAGGGAAGACAGAAAAAATAAAAAA 2636  
Db 2639 TAAAGTGTGGGAGAGACAAAAAATAAAAAA 2669

RESULT 2

US-09-985-799-101  
; Sequence 101, Application US/09985799  
; Patent No. RE38392  
; GENERAL INFORMATION:  
; APPLICANT: THOMPSON, Timothy C.  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES  
; NUMBER OF SEQUENCES: 175  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BAKER & BOTTS, L.L.P.  
; STREET: 1299 Pennsylvania Avenue, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004-2400  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/985,799  
; FILING DATE: 06-NOV-1995  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/594,031  
; FILING DATE: 30-JAN-1996  
; APPLICATION NUMBER: 60/006,838  
; FILING DATE: 16-NOV-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Remenick, James  
; REGISTRATION NUMBER: 36,902  
; REFERENCE/DOCKET NUMBER: 0A146-0110  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-639-7700  
; TELEFAX: 202-639-7890  
; TELEX: <Unknown>

; INFORMATION FOR SEQ ID NO: 101:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2669 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: <Unknown>  
; ORIGINAL SOURCE:  
; SEQUENCE DESCRIPTION: SEQ ID NO: 101:  
US-09-985-799-101  
  
Query Match 95.7%; Score 2547.8; DB 1; Length 2669;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 2591; Conservative 0; Mismatches 17; Indels 3; Gaps 3;  
  
QY 28 AACCTTGGTGGCTGCGTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCCTGGGAT 87  
Db 60 ACCTTGAGTGGCTGCGTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCCTGGGAT 119  
QY 88 TTCTGCTCCTGGCTGCAAGATTCACCTTGATGCCGCCAAACGATTTTCATGATGCTGG 147  
Db 120 TTCTGCTCCTGGCTGCAAGATTCACCTTGATGCCGCCAAACGATTTTCATGATGCTGG 179  
QY 148 GCAATGAAGACCTTCTCTTACATGAGGAGGACCAATCAATTAATGGCTGGCTTCTG 207  
Db 180 GCAATGAAGACCTTCTCTTACATGAGGAGGACCAATCAATTAATGGCTGGCTTCTG 239  
QY 208 ATGAAATGACTGGAATGAAAAAATCTACACCAAGTGTGGAAGCGGGAGACATGAGGTGA 267  
Db 240 ATGAAATGACTGGAATGAAAAAATCTACACCAAGTGTGGAAGCGGGAGACATGAGGTGA 299  
QY 268 AAAACTCCTGGAAGGAGGCGCTGTGCAGGCGGTCTGACCAAGTACTCACAGCCCTCG 327  
Db 300 AAAACTCCTGGAAGGAGGCGCTGTGCAGGCGGTCTGACCAAGTACTCACAGCCCTCG 359  
QY 328 TGGGCTCAATATATAACATTTGCGGTGAACTGATATTCCTAGATGCCAAAGGAAGATG 387  
Db 360 TGGGCTCAATATATAACATTTGCGGTGAACTGATATTCCTAGATGCCAAAGGAAGATG 419  
QY 388 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGCGTGTATCTCTGATC 447  
Db 420 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGCGTGTATCTCTGATC 479  
QY 448 CATATGTTTACACTGGACAGCATGCTCAGAGGACAGTGAAGGAAAAATGGCACCGGCC 507  
Db 480 CATATGTTTACACTGGACAGCATGCTCAGAGGACAGTGAAGGAAAAATGGCACCGGCC 539  
QY 508 AAAGCCATCATAACTCTTCCCTGATGGGAAACCTTTTCTCACCACCCCGGATGGAGAA 567  
Db 540 AAAGCCATCATAACTCTTCCCTGATGGGAAACCTTTTCTCACCACCCCGGATGGAGAA 599  
QY 568 GATGGAATTTATCTACGTTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGACGAT 627  
Db 600 GATGGAATTTATCTACGTTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGACGAT 659  
QY 628 GTTCAGTGAGATTTCTGTGAACACAGCCAAATGTGACACTTGGGCTCAACTCATGGAAG 687  
Db 660 GTTCAGTGAGATTTCTGTGAACACAGCCAAATGTGACACTTGGGCTCAACTCATGGAAG 719  
QY 688 TGAATGCTTACAGAAGACATGGACGGGCATATGTTTCCCATGATGTTCCAGAAAGACGAT 747  
Db 720 TGAATGCTTACAGAAGACATGGACGGGCATATGTTTCCCATGATGTTCCAGAAAGACGAT 779  
QY 748 ACGTGGTAACAGATCAGATTTCTGTTTGTGACTATGTTCCAGAAAGACGATCGAAAT 807  
Db 780 ACGTGGTAACAGATCAGATTTCTGTTTGTGACTATGTTCCAGAAAGACGATCGAAAT 839  
QY 808 CATCCGACGAAACCTTCC-CAAAGATCTCCCATTTATGTTTGTGATGCTCTGATTCATGATC 866  
Db 840 CATCCGACGAAACCTTCTCTCAAAGATCTCCCATTTATGTTTGTGATGCTCTGATTCATGATC 899

QY 867 CTAGCCACTTCCCTCAATTAATCTACCAATTAATCAAGTGGAGCTTCGGGATTAATCTG 926  
 Db 900 CTAGCCACTTCCCTCAATTAATCTACCAATTAATCAAGTGGAGCTTCGGGATTAATCTG 959  
 QY 927 GCCTGTTGTTTCCACCAATCATATCTGTAATCAACAGTATGCTCAATGGAACCTTCA 986  
 Db 960 GCCTGTTGTTTCCACCAATCATATCTGTAATCAACAGTATGCTCAATGGAACCTTCA 1019  
 QY 987 GCCTTAACCTCACTGTGAAAGCTGCAGCACCAAGACCTTGTCCGCCACCGCCACCAAC 1046  
 Db 1020 GCCTTAACCTCACTGTGAAAGCTGCAGCACCAAGACCTTGTCCGCCACCGCCACCAAC 1079  
 QY 1047 CCAGACCTTCAAAACCCACCCCTTCTTAAGACCTGCTGTGACACACCCCTGGAGCTGA 1106  
 Db 1080 CCAGACCTTCAAAACCCACCCCTTCTTAAGACCTGCTGTGACACACCCCTGGAGCTGA 1139  
 QY 1107 GTAGGATTCCTGATGAAGCTGCCAGATTAAAGATATGCCACCTTCAAGCCACCATCA 1166  
 Db 1140 GTAGGATTCCTGATGAAGCTGCCAGATTAAAGATATGCCACCTTCAAGCCACCATCA 1199  
 QY 1167 CAATTTGTAGGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGCTTCCTGATGCCGG 1226  
 Db 1200 CAATTTGTAGGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGCTTCCTGATGCCGG 1259  
 QY 1227 TGCCATGGCTGAAAGCTCCCTAAATAGACTTTGTGTGTGACCTGCCAAGGGAGATCCCA 1286  
 Db 1260 TGCCATGGCTGAAAGCTCCCTAAATAGACTTTGTGTGTGACCTGCCAAGGGAGATCCCA 1319  
 QY 1287 CGGAGTCTGTACCATCACTTCTGACCCACCTGGAGATCAACAGACACAGCTGTCA 1346  
 Db 1320 CGGAGTCTGTACCATCACTTCTGACCCACCTGGAGATCAACAGACACAGCTGTCA 1379  
 QY 1347 GCCTGTGATGTGGATGAGTGTCTGTCTGTGACTGTGAGACAACTTCAATGGTCTG 1406  
 Db 1380 GCCTGTGATGTGGATGAGTGTCTGTCTGTGACTGTGAGACAACTTCAATGGTCTG 1439  
 QY 1407 GGAAGTCTGTGAACTTCACTGAGGATGACACAGCTTGGCTTCAAGGACACCC 1466  
 Db 1440 GGAAGTCTGTGAACTTCACTGAGGATGACACAGCTTGGCTTCAAGGACACCC 1499  
 QY 1467 TGATTTCTGTTCTGACAGAGACCCAGCTTCTTAAAGATGGCAACAGTGCCTGTA 1526  
 Db 1500 TGATTTCTGTTCTGACAGAGACCCAGCTTCTTAAAGATGGCAACAGTGCCTGTA 1559  
 QY 1527 TCTCCGTTGGCTGTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAAAC 1586  
 Db 1560 TCTCCGTTGGCTGTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAAAC 1619  
 QY 1587 ACAAGGAATACACCCCAATAGAAATAGTCTGGGATGTGTCAGAACCAAGGCTGA 1646  
 Db 1620 ACAAGGAATACACCCCAATAGAAATAGTCTGGGATGTGTCAGAACCAAGGCTGA 1679  
 QY 1647 GTGCTTTCTCAACCGTGAAGACCGTGTCTTCCCGGAAACCCAGGAAAGATCCGC 1706  
 Db 1680 GTGCTTTCTCAACCGTGAAGACCGTGTCTTCCCGGAAACCCAGGAAAGATCCGC 1739  
 QY 1707 TACTCAAAACCAAGAAATTTAAAGGAGTTTCTTAAATTTTCGACCTTGTCTGAAGTCA 1766  
 Db 1740 TACTCAAAACCAAGAAATTTAAAGGAGTTTCTTAAATTTTCGACCTTGTCTGAAGTCA 1799  
 QY 1767 CTTTTCAGTGCATGTGATGTGAGATGTGCTGGAGTGCCTTAAACCTTTTTCCTAAAG 1826  
 Db 1800 CTTTTCAGTGCATGTGATGTGAGATGTGCTGGAGTGCCTTAAACCTTTTTCCTAAAG 1859  
 QY 1827 ATTAATGTTAAATAGATATGTTGTTTGGGAAAGTTGAATTTTATAGGTTAAATGTCA 1886  
 Db 1860 ATTAATGTTAAATAGATATGTTGTTTGGGAAAGTTGAATTTTATAGGTTAAATGTCA 1919  
 QY 1887 TTTTAGAGATGGGAGAGGGATTAATATCTGACAGGAGCTTTCAGCAATGTTTGAACCTGAT 1946  
 Db 1920 TTTTAGAGATGGGAGAGGGATTAATATCTGACAGGAGCTTTCAGCAATGTTTGAACCTGAT 1979

QY 1947 AAAAGCAACTTAGCAAGGCTTCTTTTCAATATTTTATGTTTCACTTATTAAGCTTAG 2006  
 Db 1980 AAAAGCAACTTAGCAAGGCTTCTTTTCAATATTTTATGTTTCACTTATTAAGCTTAG 2039  
 QY 2007 GTAACCTAGTAGGATAGAAACACTGTGTCGAGAGTAAAGGAGAGAAAGTACTATTGATTA 2066  
 Db 2040 GTAACCTAGTAGGATAGAAACACTGTGTCGAGAGTAAAGGAGAGAAAGTACTATTGATTA 2099  
 QY 2067 GAGCTTAACCCAGGTTAACTGCAAGAGAGAGCGGAGTACTTTTCAAGCTTTCATGTAAC 2126  
 Db 2100 GAGCTTAACCCAGGTTAACTGCAAGAGAGAGCGGAGTACTTTTCAAGCTTTCATGTAAC 2159  
 QY 2127 TATGCATAAAGCCAAATAGTCCAGTTTCTAAGATCATGTTTCCAAAGCTTAATGTAAC 2186  
 Db 2160 TATGCATAAAGCCAAATAGTCCAGTTTCTAAGATCATGTTTCCAAAGCTTAATGTAAC 2219  
 QY 2187 CTTCAATACACACTCATGAACCTCCTGATGGAACATAACAGCCCAAGCTTGTGATGA 2246  
 Db 2220 CTTCAATACACACTCATGAACCTCCTGATGGAACATAACAGCCCAAGCTTGTGATGA 2279  
 QY 2247 TGTGCACACTTGTAGACTCAGAAAAAATACCTCTCTCATAAATGGTGGGAGTATTTG 2306  
 Db 2280 TGTGCACACTTGTAGACTCAGAAAAAATACCTCTCTCATAAATGGTGGGAGTATTTG 2339  
 QY 2307 GTGACAACTACTTGTGCTGAGTGAAGAAATGATATTCATATATTCATTTATTTCCA 2366  
 Db 2340 GTGACAACTACTTGTGCTGAGTGAAGAAATGATATTCATATATTCATTTATTTCCA 2399  
 QY 2367 TGGACATTTAGTGTGCTTTTATATACCGGATGATGCTGAGTGAACCTCTTGTGTA 2426  
 Db 2400 TGGACATTTAGTGTGCTTTTATATACCGGATGATGCTGAGTGAACCTCTTGTGTA 2459  
 QY 2427 TATTTTCAAAATTTTGTATAGTGGTGCACATATTTGAAATCAAAATATTAAGACTTTCC 2486  
 Db 2460 TATTTTCAAAATTTTGTATAGTGGTGCACATATTTGAAATCAAAATATTAAGACTTTCC 2518  
 QY 2487 AAAAAATTTGGTCCCTGGTTTTTTCATGGCAACTGATCAGTAAGGATTTCCCTCTGTTG 2546  
 Db 2519 AAAAAATTTGGTCCCTGGTTTTTTCATGGCAACTGATCAGTAAGGATTTCCCTCTGTTG 2578  
 QY 2547 GAACTTAAACCATTTACTATATGTTAGACAGACATTTTTTTTTTTTCTTCTCTGAAAAA 2606  
 Db 2579 TAACTTAAACCATTTACTATATGTTAGACATGACATTTTCTCTCTCTCTGAAAAA 2638  
 QY 2607 -AAAAATGAGGAGAGACAAAAA 2636  
 Db 2639 TAAAGTGTGGAGAGACAAAAA 2669

RESULT 3  
 US-08-594-031-91  
 ; Sequence 91, Application US/08594031  
 ; Patent No. 5783182  
 ; GENERAL INFORMATION:  
 ; APPLICANT: THOMPSON, Timothy C.  
 ; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES  
 ; NUMBER OF SEQUENCES: 175  
 ; CORRESPONDENCE ADDRESSES:  
 ; ADDRESSEE: BAKER & BORTS, L.L.P.  
 ; STREET: 1299 Pennsylvania Avenue, N.W.  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20004-2400  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FASTSEQ Version 1.5  
 ; CURRENT APPLICATION NUMBER: US/08/594,031  
 ; FILING DATE: 30-JAN-1996  
 ; CLASSIFICATION: 435

PRIOR APPLICATION DATA: APPLICATION NUMBER: 60/006,838  
FILING DATE: 16-NOV-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Remenick, James  
REGISTRATION NUMBER: 36,902  
REFERENCE/DOCKET NUMBER: 0A146-0110  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-639-7700  
TELEFAX: 202-639-7890  
TELEX:

```

: INFORMATION FOR SEQ ID NO: 91:
:
: SEQUENCE CHARACTERISTICS:
:   LENGTH: 2669 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: single
:   TOPOLOGY: linear
:   MOLECULE TYPE: cDNA
:   HYPOTHETICAL: NO
:   ANTI-SENSE: NO
:   FRAGMENT TYPE:
:   ORIGINAL SOURCE:
: US-08-594-031-91

```

Query Match 95.7%; Score 2547.8; DB 1; Length 2669;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 2591; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

	Qy	28	AACCTTGGTGCTCGTCGCGGTGAGAAATTCAGCATGGGAATGCTCTACTATTTCCTGGAT	87
	Db	60	ACCTTGAGTGCCTCGCTCGGTGAGAAATTCAGCATGGGAATGCTCTACTATTTCCTGGAT	119
	Qy	88	TTCTGCTCCTGGCTGCAAGATTGCCAATTGATGCCGCCAAACGATTTCATGATGCTCG	147
	Db	120	TTCTGCTCCTGGCTGCAAGATTGCCAATTGATGCCGCCAAACGATTTCATGATGCTCG	179
	Qy	148	GCAATGAAGAACCCTTCTGCTTACATGAGGAGACAATCAATTAATGGCTGGTCTTCTG	207
	Db	180	GCAATGAAGAACCCTTCTGCTTACATGAGGAGACAATCAATTAATGGCTGGTCTTCTG	239
	Qy	208	ATGAAATGACTGGAATGA AAAA CT TACCCAGTGTGGAAGCGGGGAGACATGAGGTGGA	267
	Db	240	ATGAAATGACTGGAATGA AAAA CT TACCCAGTGTGGAAGCGGGGAGACATGAGGTGGA	299
	Qy	268	AAAACTCTGGAAGGAGGCGGTGTGAGCGCGTCTGACAGTGA CT CACAGCCCTCG	327
	Db	300	AAAACTCTGGAAGGAGGCGGTGTGAGCGCGTCTGACAGTGA CT CACAGCCCTCG	359
	Qy	328	TGSGCTCAAATAAACA TT T GCGGTGAA C CTGTATAT TCCTAGATGCCAAAAGAAATG	387
	Db	360	TGSGCTCAAATAAACA TT T GCGGTGAA C CTGTATAT TCCTAGATGCCAAAAGAAATG	419
	Qy	388	CCAATGGCAACATAGTCTATGAGAA GA A CT GCAGAAATGAGCGCTGTTATCTGCTGATC	447
	Db	420	CCAATGGCAACATAGTCTATGAGAA GA A CT GCAGAAATGAGCGCTGTTATCTGCTGATC	479
	Qy	448	CATATGTTTACA ACTGGA C AGA NTGT CAGAGACAGTGA CGGGGAAAATGGCACCGGCC	507
	Db	480	CATATGTTTACA ACTGGA C AGA NTGT CAGAGACAGTGA CGGGGAAAATGGCACCGGCC	539
	Qy	508	AAAGCCATCAATAACGTC TT CCCTGATGGGAA ACCTTTCTCACACCCCGGATGAGAA	567
	Db	540	AAAGCCATCAATAACGTC TT CCCTGATGGGAA ACCTTTCTCACACCCCGGATGAGAA	599
	Qy	568	GATGGAATTTCA TCTACGCTTCCACACACTTGGTCAGTAT TTCAGAAAATTTGGGACCAT	627
	Db	600	GATGGAATTTCA TCTACGCTTCCACACACTTGGTCAGTAT TTCAGAAAATTTGGGACCAT	659
	Qy	628	GTTTCAGTGA GAGTTTCTGTGMA CACAGCCA NTGTGACACTTGGGCCCTCAACTCATGGAA G	687
	Db	660	GTTTCAGTGA GAGTTTCTGTGMA CACAGCCA NTGTGACACTTGGGCCCTCAACTCATGGAA G	719
	Qy	688	TGACTGCTTACAGAGACATGGACGGGCATAT GTTCCCATCGCA CAAGTGA AAGATGTGT	747

Mon Sep 13 10:28:12 2004

us-10-039-272-1.rni

1800 CTTTTAGTGCATATGATGATGCTGAGTGGCTATTAACCTTTTCTCTAAG 1859  
1827 ATTATTGTTAAATAGATATGTTGGTTGGGAAGTTGAATTTTATAGGTTAAATGTCA 1886  
1860 ATTATTGTTAAATAGATATGTTGGTTGGGAAGTTGAATTTTATAGGTTAAATGTCA 1919  
1887 TTTTAGAGATGGGAGGAGGATTATACGAGGAGCTTCAGCCATGTTGTGAAACTGAT 1946  
1920 TTTTAGAGATGGGAGGAGGATTATACGAGGAGCTTCAGCCATGTTGTGAAACTGAT 1979  
1947 AAAAGCAACTTAGCAAGGCTCTTTTCAATTTATTTTATGTTTCACTTATAAGTCTTAG 2006  
1980 AAAAGCAACTTAGCAAGGCTCTTTTCAATTTATTTTATGTTTCACTTATAAGTCTTAG 2039  
2007 GTAACTAGTAGATAGAAACACTGTGTCGAGAGTAAGGAGAGAGCTACTATTGATTA 2066  
2040 GTAACTAGTAGATAGAAACACTGTGTCGAGAGTAAGGAGAGAGCTACTATTGATTA 2099  
2067 GAGCCTAACCCAGGTTAACTGCAAGAGAGGCGGATATTTTCACTTTCCATGTAACGTG 2126  
2100 GAGCCTAACCCAGGTTAACTGCAAGAGAGGCGGATATTTTCACTTTCCATGTAACGTG 2159  
2127 TATGCAATAAGCAATAGTGTAGTTCAGTTCCTAAGATCATGTTCCAAAGCTAACTGAATCCCA 2186  
2160 TATGCAATAAGCAATAGTGTAGTTCAGTTCCTAAGATCATGTTCCAAAGCTAACTGAATCCCA 2219  
2187 CTTCAATACACTCATGAACTCTGATGGAACAATAAAGGCGCCAGGCTGTGATGA 2246  
2220 CTTCAATACACTCATGAACTCTGATGGAACAATAAAGGCGCCAGGCTGTGATGA 2279  
2247 TGTGCACTCTGCTAGACTCAGAAAAAATCTACTCTCATAAATGGTGGGAGTATTTTG 2306  
2280 TGTGCACTCTGCTAGACTCAGAAAAAATCTACTCTCATAAATGGTGGGAGTATTTTG 2339  
2307 GTGCAAACTACTTTCTGCTGCTGAGTGAAGGAATGATATTCATATATTATTTATCCCA 2366  
2340 GTGCAAACTACTTTCTGCTGCTGAGTGAAGGAATGATATTCATATATTATTTATCCCA 2399  
2367 TGGACATTTAGTTAGTGTCTTTTATATACGAGCATGCTGAGTGACACTCTTGTA 2426  
2400 TGGACATTTAGTTAGTGTCTTTTATATACGAGCATGCTGAGTGACACTCTTGTA 2459  
2427 TATTTCCAAATTTTGTATAGTGCCTGCACATATTTGAATCAAAATATTAAGACTTTCC 2486  
2460 TATTTCCAAATTTTGTATAGTGCCTGCACATATTTGAATCAAAATATTAAGACTTTCC 2518  
2487 AAAAATTTGGTCCCTGCTGTTTTCATGGCAACTTGTACAGTAAGGATTTCCCTCTGTTG 2546  
2519 AAAGATGAGTCCCTGCTGTTTTCATGGCAACTTGTACAGTAAGGATTTCCCTCTGTTG 2578  
2547 GAACATAAACCATTTACTATATGTTAGACAGACATTTTTTTTTTCTCTCTCTGAAAAA 2606  
2579 TAACTAAACCATTTACTATATGTTAGACATGACATTTCTTTTCTCTCTCTCTCTGAAAAA 2638  
2607 -AAANTGAGGAGAGACAAAAA 2636  
2639 TAAAGTGTGGAAGAGACAAAAA 2669

RESULT 4  
US-08-594-031-101  
; Sequence 101, Application US/08594031  
; Patent No. 5783182  
; GENERAL INFORMATION:  
; APPLICANT: THOMPSON, Timothy C.  
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES  
; NUMBER OF SEQUENCES: 175  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BAKER & BOTT, L.L.P.  
; STREET: 1299 Pennsylvania Avenue, N.W.  
; City: Washington  
; STATE: DC  
; COUNTRY: USA

ZIP: 20004-2400  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/594.031  
FILING DATE: 30-JAN-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/006,838  
FILING DATE: 16-NOV-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Remenick, James  
REGISTRATION NUMBER: 36,902  
REFERENCE/DOCKET NUMBER: 0A146-0110  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-639-7700  
TELEFAX: 202-639-7890  
TELEX:  
INFORMATION FOR SEQ ID NO: 101:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2669 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLSCULE TYPE: cdna  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
US-08-594-031-101  
Query Match 95.7%; Score 2547.8; DB 1; Length 2669;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 2591; Conservative 0; Mismatches 17; Indels 3; Gaps 3;  
QY 28 AACCTTGGTGCCTGCTCGTGGAGATTCAGATGGAATGTCTCTACTATTTCTCTGGAT 87  
DB 60 ACCTTGAGTGCCTGCTCGTGGAGATTCAGATGGAATGTCTCTACTATTTCTCTGGAT 119  
QY 88 TTCTGCTCTGCTGCAAGATTGCCACTTGTATGCCGCCCAACGATTTTCATGATGCTG 147  
DB 120 TTCTGCTCTGCTGCAAGATTGCCACTTGTATGCCGCCCAACGATTTTCATGATGCTG 179  
QY 148 GCATCAAAAGACCTTCTGCTTACATGAGGAGCACAATCAATTAATGGTGTCTTCTG 207  
DB 180 GCATCAAAAGACCTTCTGCTTACATGAGGAGCACAATCAATTAATGGTGTCTTCTG 239  
QY 208 ATCAAAATGACTGGAATGAAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGA 267  
DB 240 ATCAAAATGACTGGAATGAAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGA 299  
QY 268 AAAACTCTGGAAGGAGGCGCGTGTGAGCGGTCTGACCGAGTCTGACCGAGTCTGACCG 327  
DB 300 AAAACTCTGGAAGGAGGCGCGTGTGAGCGGTCTGACCGAGTCTGACCGAGTCTGACCG 359  
QY 328 TGGGCTCAAAATATAACATTTTGGGTGAACTCTGATATTTCCCTAGATGCAAAAGGAAGATG 387  
DB 360 TGGGCTCAAAATATAACATTTTGGGTGAACTCTGATATTTCCCTAGATGCAAAAGGAAGATG 419  
QY 388 CCAATGGCAACATAGTCTATGAGAAGAACTCTGAGAAGAACTGAGAAGAACTGAGGTGAT 447  
DB 420 CCAATGGCAACATAGTCTATGAGAAGAACTCTGAGAAGAACTGAGGTGATGAGGTGAT 479  
QY 448 CATATGTTTCAACTGGACAGCATGTCAGAGACATGTCAGAGACATGACGGGAAATGGCACCG 507  
DB 480 CATATGTTTCAACTGGACAGCATGTCAGAGACATGTCAGAGACATGACGGGAAATGGCACCG 539  
QY 508 AAGCCCATCAACGCTCTCCCTGATGGAAACCTTTTCTCTCACCACCCCGATGGAGAA 567  
DB 540 AAGCCCATCAACGCTCTCCCTGATGGAAACCTTTTCTCTCACCACCCCGATGGAGAA 599



QY 568 GATGAATTTTCATCTACGTCCTTCCACACACTTGGTCAGTATTTTCAGAAATTTGGGACGAT 627  
DB 600 GATGGAATTTTCATCTACGTCCTTCCACACACTTGGTCAGTATTTTCAGAAATTTGGGACGAT 659  
QY 628 GTTCAGTGAGAGTTTCTGTGGAACACAGCCAAATGTGACACTTGGGCTCAACTCATGGAAG 687  
DB 660 GTTCAGTGAGAGTTTCTGTGGAACACAGCCAAATGTGACACTTGGGCTCAACTCATGGAAG 719  
QY 688 TGACGTCTACAGAGACATGACCGGCATATGTTCCCATCGCAACAGTGAAGATGTGT 747  
DB 720 TGACGTCTACAGAGACATGACCGGCATATGTTCCCATCGCAACAGTGAAGATGTGT 779  
QY 748 AGTGTTTAAACAGATCAGATTCCTGTGTTGTGTCACATATGTTCCAGAGACGATCGAAAT 807  
DB 780 AGTGTTTAAACAGATCAGATTCCTGTGTTGTGTCACATATGTTCCAGAGACGATCGAAAT 839  
QY 808 CATCCGAGAAACCTTCC-CAAAGATCTCCCATATATGTTTGAATGCTGTGATTCATGATC 866  
DB 840 CATCCGAGAAACCTTCCCAAAGATCTCCCATATATGTTTGAATGCTGTGATTCATGATC 899  
QY 867 CTAGCACCTCTCTCAATTTCTACCATTAACATAAGTGGAGCTTCGGGGATATATCTG 926  
DB 900 CTAGCACCTCTCTCAATTTCTACCATTAACATAAGTGGAGCTTCGGGGATATATCTG 959  
QY 927 GCCTGTTTGTTCACCAATCATCTGTGAATCACACGTATGTGCTCAATGGAACTTCA 986  
DB 960 GCCTGTTTGTTCACCAATCATCTGTGAATCACACGTATGTGCTCAATGGAACTTCA 1019  
QY 987 GCCTTAACTCACTGTGAAAGCTGCAGCACAGGACCTTGTCCGCCACCGCCACCCAC 1046  
DB 1020 GCCTTAACTCACTGTGAAAGCTGCAGCACAGGACCTTGTCCGCCACCGCCACCCAC 1079  
QY 1047 CCAGACCTTCAAACCCACCTTCTTTAGACCTGTGTGGAACAAACCCCTGGAGCTGA 1106  
DB 1080 CCAGACCTTCAAACCCACCTTCTTTAGACCTGTGTGGAACAAACCCCTGGAGCTGA 1139  
QY 1107 GTAGATTCCTGATGAAACCTGCCAGATTAACAGATATGGCCATTTTCAAGCCACCATCA 1166  
DB 1140 GTAGATTCCTGATGAAACCTGCCAGATTAACAGATATGGCCATTTTCAAGCCACCATCA 1199  
QY 1167 CAATTTAGAGGGATCTTAGAGGTTAACTATCATCTCCAGATGACAGCTCTGTATGCGG 1226  
DB 1200 CAATTTAGAGGGATCTTAGAGGTTAACTATCATCTCCAGATGACAGCTCTGTATGCGG 1259  
QY 1227 TGCCATGCGCTGAAGCTCCCTAATAGACTTTGTGTGACCTGCAAGGGAGCAATCCCA 1286  
DB 1260 TGCCATGCGCTGAAGCTCCCTAATAGACTTTGTGTGACCTGCAAGGGAGCAATCCCA 1319  
QY 1287 CGGAGGTCTGTACCATCAATTTCTGACCCCTCGAGATCACAGCTCTCAATGGGTCTG 1346  
DB 1320 CGGAGGTCTGTACCATCAATTTCTGACCCCTCGAGATCACAGCTCTCAATGGGTCTG 1379  
QY 1347 GCCTGTGGATGGATGAGATGTGTCTGTGACTGTGAGACGAACCTTCAATGGGTCTG 1406  
DB 1380 GCCTGTGGATGGATGAGATGTGTCTGTGACTGTGAGACGAACCTTCAATGGGTCTG 1439  
QY 1407 GGAAGTCTGTGAACTTCACTTGGGGATGACAAAGCTGGCTCTCAAGAGACCC 1466  
DB 1440 GGAAGTCTGTGAACTTCACTTGGGGATGACAAAGCTGGCTCTCAAGAGACCC 1499  
QY 1467 TGAATTTCTGTCTGACAGAGACCCAGCTCGCTTTAAGATGGCAACAGTGCCTGA 1526  
DB 1500 TGAATTTCTGTCTGACAGAGACCCAGCTCGCTTTAAGATGGCAACAGTGCCTGA 1559  
QY 1527 TCTCCGTGGCTGTGGCCATATTTGTCACTGTGATCTCCCTTGTGCTGTACAAAAC 1586  
DB 1560 TCTCCGTGGCTGTGGCCATATTTGTCACTGTGATCTCCCTTGTGCTGTACAAAAC 1619  
QY 1587 ACAGGATACACCCCAATGAAATAGTCTGGGAATGTGGTCAAGCAAGAGGCTGA 1646  
DB 1620 ACAAGGAATACACCCCAATGAAATAGTCTGGGAATGTGGTCAAGCAAGAGGCTGA 1679

QY 1647 GTGTCTTTCTAAACCGTGCAAAAGCGGTGTTCTTCCGGGAAACACAGGAAAGATCCCG 1706  
DB 1680 GTGTCTTTCTAAACCGTGCAAAAGCGGTGTTCTTCCGGGAAACACAGGAAAGATCCCG 1739  
QY 1707 TACTCAAAACCAAGAAATTAAGAGATTTCTTAATTTTCGACCTGTGTTCTGAAGCTCA 1766  
DB 1740 TACTCAAAACCAAGAAATTAAGAGATTTCTTAATTTTCGACCTGTGTTCTGAAGCTCA 1799  
QY 1767 CTTTTCAGTGCCATTAATGATGTGTGAGATGTGTGAGTGGCTTAAACCTTTTCTTAAAG 1826  
DB 1800 CTTTTCAGTGCCATTAATGATGTGTGAGATGTGTGAGTGGCTTAAACCTTTTCTTAAAG 1859  
QY 1827 ATTATGTTAAATAGATATTTGTGTTTGGGAAAGTTGAATTTTATAGTTTAAATGTCA 1886  
DB 1860 ATTATGTTAAATAGATATTTGTGTTTGGGAAAGTTGAATTTTATAGTTTAAATGTCA 1919  
QY 1887 TTTTATAGATGGGAGAGGATTAATGATGAGTGGCTTAAACCTTTTCACTTATAAGTCTAG 1946  
DB 1920 TTTTATAGATGGGAGAGGATTAATGATGAGTGGCTTAAACCTTTTCACTTATAAGTCTAG 1979  
QY 1947 AAAAGCAACTTAGCAAGGCTTCTTTTCAATTTTATTTTATGTTTCACTTATAAGTCTAG 2006  
DB 1980 AAAAGCAACTTAGCAAGGCTTCTTTTCAATTTTATTTTATGTTTCACTTATAAGTCTAG 2039  
QY 2007 GTAACTAGTAGATAGAAACACTGTGTCTCCGAGAGTAAAGAGAGACTACTATTGATTA 2066  
DB 2040 GTAACTAGTAGATAGAAACACTGTGTCTCCGAGAGTAAAGAGAGACTACTATTGATTA 2099  
QY 2067 GAGCCTTAACCCAGGTTAACTGCAAGAGAGCGGGATACCTTTCAGCTTCCATGTAATG 2126  
DB 2100 GAGCCTTAACCCAGGTTAACTGCAAGAGAGCGGGATACCTTTCAGCTTCCATGTAATG 2159  
QY 2127 TATCATATAAGCCAAATGTAGTCCAGTTTCTAAGATCATGTTTCCAGCTTAACTGAATCCCA 2186  
DB 2160 TATCATATAAGCCAAATGTAGTCCAGTTTCTAAGATCATGTTTCCAGCTTAACTGAATCCCA 2219  
QY 2187 CTTCAATACACACTCATGAACTCTGTATGGAACAAATAACAGGCCCAAGCTGTGGTATGA 2246  
DB 2220 CTTCAATACACACTCATGAACTCTGTATGGAACAAATAACAGGCCCAAGCTGTGGTATGA 2279  
QY 2247 TGTGCACTTGTCTAGACTCAGAAATAATCTACTCTATAAATGGGTGGGAGTATTTG 2306  
DB 2280 TGTGCACTTGTCTAGACTCAGAAATAATCTACTCTATAAATGGGTGGGAGTATTTG 2339  
QY 2307 GTGCAACCTACTTGTGCTGCTGAGTGAAGAAATGATATTCATATATTCATTTATTTCCA 2366  
DB 2340 GTGCAACCTACTTGTGCTGCTGAGTGAAGAAATGATATTCATATATTCATTTATTTCCA 2399  
QY 2367 TGGACATTTAGTTAGTGTCTTTTATATACAGGCATGATGCTGAGTGACACTCTTGTGTA 2426  
DB 2400 TGGACATTTAGTTAGTGTCTTTTATATACAGGCATGATGCTGAGTGACACTCTTGTGTA 2459  
QY 2427 TATTTCCAAATTTTGTATAGTGTGTCACATATTTGAAATCAAAATATTAAAGACTTTCC 2486  
DB 2460 TATTTCCAAATTTTGTATAGTGTGTCACATATTTGAAATC-ATATATTAAAGACTTTCC 2518  
QY 2487 AAAAATTTGTCTCCCTGTTTTCATGCAACTTGTATGAGTGAAGATTTCCCTCTGTTTG 2546  
DB 2519 AAAGATGAGGTCCCTGTTTTCATGCAACTTGTATGAGTGAAGATTTCCCTCTGTTTG 2578  
QY 2547 GAACTAAAACCAATTTACTATATGTTAGACAAGACATTTTCTTCTTCTTCTTCTGAAAAA 2606  
DB 2579 TAACTAAAACCAATTTACTATATGTTAGACATGACATTTCTTCTTCTTCTTCTGAAAAA 2638  
QY 2607 -AAAAATGAGGAGAGACAAAAA 2636  
DB 2639 TAAAGTGTGGAAGAGACAAAAA 2669

RESULT 5

US-09-197-970B-4

; Sequence 4, Application US/09197970B

; Patent No. 6664385

GENERAL INFORMATION:  
APPLICANT: Michele Sanicola-Nadel  
Joseph V. Bonventre  
Catherine A. Hession  
Takaharu Ichimura  
Henry Wei  
Richard L. Cate  
TITLE OF INVENTION: MODULATORS OF TISSUE REGENERATION  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Biogen, Inc.  
STREET: 14 Cambridge Center  
CITY: Cambridge  
STATE: MA  
COUNTRY: USA  
ZIP: 02142  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/197,970B  
FILING DATE: 23-NO. 6664385-1998  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/018,228  
FILING DATE: 24-MAY-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Levine, Leslie M.  
REGISTRATION NUMBER: 35,245  
REFERENCE/DOCKET NUMBER: A010 PCT CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 679-2810  
TELEFAX: (617) 679-2838  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2303 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cdna  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 107..1822  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-197-970B-4

Query Match 39.5%; Score 1051.2; DB 4; Length 2303;  
Best Local Similarity 74.1%; Pred. No. 1.3e-216;  
Matches 1405; Conservative 0; Mismatches 443; Indels 48; Gaps 4;  
QY 31 CTTGGTGCCTGCGTGAGAAATCAGCATGGAATGTCTCTACTATTTCCTGGGATTC 90  
Db 78 CTGGCTGAGGCGGGCGCTCCGAGTCAGCATGGAAGTCTCTCGGGGCTCTGGTATTC 137  
QY 91 TGTCTCTGGCTGAAGATTGCCATTCGATGCCCCCAACGATTTGATGCTCTGGGA 150  
Db 138 TGTCTGGCTGAGGACTGCCCTCCAGCGGCCCAAGCGGTTCCTGGTATGCTGGGCC 197  
QY 151 ATCAAGACCTTCTGCTTACATGAGGGAGCACAAATTAATTAATGGCTGCTTCTGATG 210  
Db 198 ATGAGCAGTATCCGATCAGATGAGGGAGAACCAATTAATGCTGGTCTTTCAGATG 257  
QY 211 AAAATGATGGAATGAAAACTTACCCAGTGTGGAAGCGGGGAGAGCATGAGTGGAAA 270  
Db 258 AAAATGAATGGATGAACAGCTGTATCCAGTGTGAGGAGGGGAGGAGGAGTGAAGG 317  
QY 271 ACTCTGGAAGGAGGCGGTGTGAGCGGTCTGACCACTGACTCACAGCCCTCGTGG 330  
Db 318 ACTCCTGGGAGGAGGCGGTGTGAGCGAGCCCTAACCGATGATTACCGGCTTGTGG 377  
QY 331 GCTCAATATAACATTTTGGGTGAACCTGTATATTCCTAGATGCCAAAAGGAAGTCCA 390

Db 378 GTTCCAAATATCACCTTCGTAGTGAACCTGGTTCCTCCAGATGCCAAGGAAGATGCCA 437  
QY 391 ATGCGCAATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTCTGATCCAT 450  
Db 438 ACGCAATATGCTCTATGAGAGGAACCTGCAAGTGTATTTGGAGCTGGCTTCTGACCGGT 497  
QY 451 ATGTTTACAACCTGACAGCATGCTCAGAGGACAGTGCAGGGGAAAAATGGCCCGCCAAA 510  
Db 498 ATGCTACAACCTGACACAGGGGACAGATGAGGACTGGGAAGACAGACACAGCCCAAG 557  
QY 511 GCCATCATAGCTTTCCTGATGGGAAACCTTTTCTCCACACCCCGGATGGAGAAGAT 570  
Db 558 GCCAGCACCTCAGGTTCCTCCGACGGGAAGCCCTTCCTCGCCCCACCGACGGAAGAAAT 617  
QY 571 GGAATTTTCATCTACGTCTTCCACACACTTGTGTGTCAGTATTTCCAGAAAATTTGGACGATGTT 630  
Db 618 GGAATTTTCATCTACGTCTTCCACACACTTGTGTGTCAGTATTTCCAGAAAATTTGGACGATGTT 677  
QY 631 CAGTGAGAGTTTCTGTGAACACAGCCAAATGTGACACTTTGGGCTTCAACTCATGTGAAGTGA 690  
Db 678 CAGCACGAGTTTCTATAAAACACAGTCAACTTGACAGTTTGGGCTCAGGTCATGTGAAGTGA 737  
QY 691 CTGCTACAGAGACATGGAAGGACATATGTTCCCATCGCAAGTGAAGATGTGTAGC 750  
Db 738 TTGTCTTTGGAAGACACGGCCGGGCAATATTTCCCATCTTCCAAAGTGAAAGACGTGTATG 797  
QY 751 TGGTAACAGATCAGATTCCTGTGTTTGTGATCTATTTCCAGAAAGCAAGTCAATTCAT 810  
Db 798 TGATAACAGATCAGATTCCTGTATTTCTGTCACCAATGTAACAGAGAAATGACCGAACTCT 857  
QY 811 CCGACGAAACCTTCC-CAAAGATCTCCCAATATATTTGATGTCTCTGATTCATGATCCTA 869  
Db 858 CTGATGAAACCTTCTCTCAGAGACCTCCCAATTTCTTCGATGTCTCTCATTTACGATCCCA 917  
QY 870 GCCACTTCTCAATTTATTTCTACCACTTAACAGTGGAGCTTCGGGGATATATCTTGGCC 929  
Db 918 GTCAATTTCTCAACTACTCTGCCATTTCTCAAGTGGAACTTTGGGGACAACTTGGGCC 977  
QY 930 TGTGTTGTTTCCACCAATCATACTGTGAATCAACAGTATGTCTCAATGGAACCTTCAAGCC 989  
Db 978 TGTGTTGTTTCCCAACATCAGCTTTGAATCAGCATGTGTCTCAATGGAACCTTCAACT 1037  
QY 990 TTAACCTCACTGTGAAAGCTGCAGACCAAGGACCTTGTGCGGCCACCGCCACAC- 1043  
Db 1038 TTAACCTCACTGTGAAAGCTGCAGTGCAGGACCACTGCGCCCTCACCCACACCTTCGCCCT 1097  
QY 1044 -----CACCCAGACCTTCAAAACCCACCC 1067  
Db 1098 CTTCTTCGACTTCTCCTTCGCTGCAATCTTCGCTTCCCCACATTTATCAACACTGATC 1157  
QY 1068 CTTCTTTAGGACTGCTGCTGTGACAAACCCCTCGAGCTGAGTAGGATTCCTGATGAAACT 1127  
Db 1158 CTTCTTTAATGCTACTGSCCAATCCATGAGCTGAGTGACATTTCCAATGAAACT 1217  
QY 1128 GCCAGATTAAAGATATGCCCCTTTCAAGCCCAACATCAATTTAGAGGAACTCTTAG 1187  
Db 1218 GCCGATAAAGATATGTTTACTTTCAGAGCCACCAATCAATTTAGATGGAATCCTAG 1277  
QY 1188 AGGTTAAACATCATCCAGATGACAGAGCTCTGATGCGGGTGCCTGCGCTGAAAGCTCCC 1247  
Db 1278 AAGTCAACATCATCCAGGAGGAGATGTTCCCAATCCCAACCGGAGCTGACACTCAC 1337  
QY 1248 TAATAGACTTTGTGCTGAGCTGCAAGGGAGGACATTCCTCCAGGAGGCTGTGTACCATTT 1307  
Db 1338 TGATGACTTCATTTGATGCTGCAAGGGGCCACTCCCAAGGAGCTGTACGATCATCT 1397  
QY 1308 CTGACCCCACTCGAGATCAACCCAGAACACAGTCTGACCCCTTGTGGATGTGGATGAGA 1367  
Db 1398 CTGACCCCACTCGAGATCGCCCAAGAACAGAGGTGTGACGCCCGGTGGCTGTGGATGAGC 1457  
QY 1368 TGTGCTGTGCTGATGTGAGACGAACCTTCAATGGTCTGGGAGCTACTGTGTGAACCTCA 1427

Db 1458 TGTGCTCTCTGTCCTGTCAGGAGAGCCTTCAATGGTCCGGACAGTACTGTGTAATTTCA 1517  
QY 1428 CCTGGGGATGACACAAGCTCGCTCTCAGGAGCACCCTGATTCTGTTCCTGACAGAG 1487  
Db 1518 CTCTGGGAGCAGATGCAAGCTGCGCTTACCAGCGCCTGTATCTATCCCTGGCAAG 1577  
QY 1488 ACCAGCCTGCTTTAAGGATGGCAACAGTGCCTGATCTCCGTTGGCTGCTTGGCCA 1547  
Db 1578 ACCTAGGCTCCCTCTGAGAACAGTGAATGGTGTCTGTATCTCCATTTGGCTGCTGGCCA 1637  
QY 1548 TATTCTGCTGATGATCTCTCTCTGCTGTGTAACAAACACACAGGAATACACCCCAATAG 1607  
Db 1638 TGTTGCTACCAAGTGTACCATCTCTGTGTACAAAAACACACAGACGTACAGCCCAATAG 1697  
QY 1608 AAAATAGTCTCGGAAATGGTTCAGAACAAAGCGCTGAGTCTCTCTCAACCGTGCAC 1667  
Db 1698 GAAACTGCACAGGACGTGGTCAAGGGCAAGGCGCTGAGTGTCTTCTCAGCCATGCAC 1757  
QY 1668 AAGCGTGTCTTCCGGGAAACACAGGAAAGGATCCGCTACTCAAAAAACCAAGAAATTTA 1727  
Db 1758 AAGCGCGTCTCTCCGAGGAGACCGGAGAGGATCCACTGCTCCAGGACAGCCATGGA 1817  
QY 1728 AAGGAGTTCTTAAATTCGACCTTGTCTGAAGTCTACTTTTCAGTCCCATGATGTG 1787  
Db 1818 TGCTCTAGTCTTCACTCTCACTCTGACTGGGAACCCACTCTCTGTG-CATGTATGTG 1876  
QY 1788 AGATGCTGTGAG-----TGGCTATTAACTTTTCTTAAAGATTATGTTAAATAGAT 1843  
Db 1877 ACCTGTGCAGAGTACATGACTGTAGCTGTGTTTCTTACGGATTATGTTAAATGTAT 1936  
QY 1844 ATTGTGTTTGGGAAAGTGAATTTTATAGTTTA 1879  
Db 1937 ATCAGTTTATAGGAGCGTAGTTAATTGGCAATTTTA 1972

## RESULT 6

US-09-383-586-27

; Sequence 27, Application US/09383586

; Patent No. 6242419

; GENERAL INFORMATION:

; APPLICANT: Strachan, Lorna

; APPLICANT: Sleeman, Matthew

; APPLICANT: Abernethy, Nevin

; APPLICANT: Onrust, Rene

; APPLICANT: Kumble, Anand

; APPLICANT: Muxilon, Greg

; TITLE OF INVENTION: Compounds isolated from stromal cells

; FILE REFERENCE: 11000.1037c1

; CURRENT APPLICATION NUMBER: US/09/383,586

; NUMBER OF SEQ ID NOS: 38

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 27

; LENGTH: 2213

; TYPE: DNA

; ORGANISM: Mouse

; US-09-383-586-27

Query Match 39.4%; Score 1047.8; DB 3; Length 2213;

Best Local Similarity 74.1%; Pred. No. 6,7e-216;

Matches 1404; Conservative 0; Mismatches 437; Indels 54; Gaps 4;

QY 40 TGGCTCGCTGAGATTCAGATGGAATGCTCTACTATTTCCTGGGATTTCTGCTCTGG 99

Db 3 TCGAGGCGCTCGAGTCAGATGGAAGTCTCTGCGGGTCTGGGATTTCTGCTGCTGG 62

QY 100 CTCGAAGATTGCCACTTGTATGCGGCCAAAGATTTCAATGATGCTGCGCAATGAAGAC 159

Db 63 CTCGAGACTGCTCTCCAGCTGCCAAGGATTTCTGTAATGCTGCGGCATGAACAGT 122

QY 160 CTTCTGCTTACATGAGGAGACAAATCAATTAATGCTGCTCTCTGATGAAATGACT 219

Db 123 ATCCCAATCAGATGAGAGAGCAACAACAAATTACGTGGTGGTCTTGGATGAAATGAAT 182  
QY 220 GGAATGAAAAAATCTTACCCAGTGTGGAAGCGGGGAGACATGAGTGGTGAATAAATCTCTGGA 279  
Db 183 GGGATGAAACCTTGATTCAGTGTGAGAGGGGAGACGGCAGGTGAAGGATCTCTGGG 242  
QY 280 AGGAGGCGGTGTGAGGCGGTCTTGACCAAGTGAATCACCAGCCCTCTGTTGGGCTCAATA 339  
Db 243 AAGGAGGCGGTGTGAGGCGGTCTTGACCAAGTGAATCACCAGCCCTCTGTTGGGCTCAATA 302  
QY 340 TAAATTTTGGGTGACCTGATATTTCCCTAGATGCAAAAGGAAGATGCAATGGAACA 399  
Db 303 TACCTTTTGGGTGAACTCTGGTGTTCCTCCAGATGCGAAGGAAGATGCTAATGGCAATA 362  
QY 400 TAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTCTGATCCATATGTTTACA 459  
Db 363 TCGTCTATGAGAAGAACTGCAGGAATGATTTGGGACTGACCTTGACCTGCAATGTTACA 422  
QY 460 ACTGACAGCATGGTTCAGAGGACAGTACGGGGGAAATGGCAACCGGCCAAAGCCATCATA 519  
Db 423 ACTGACTGACAGGCGCAGATGATGGTACTGGGAAGATGGCACACAGCCGACGATC 482  
QY 520 AGCTCTCTCCCTGATGGGAAACCTTTTCCCTACCAACCCCGATGGAGAGATGGAATTTCA 579  
Db 483 TCAGTTTCCCGACAGGAGGCGCTTCCCTCGCCCCCATGGATGGGAAGAAATGGAGCTTTG 542  
QY 580 TCTAGCTCTTCCACACACTTGGTCACTATTTTCCAGAAATTCGGACGATGTTCAAGTGAAG 639  
Db 543 TCTAGCTCTTCCACACACTTGGCCAGTATTTTCCAAAACCTGGGTGGTGTTCAGCACGGG 602  
QY 640 TTTCTGTGAACACAGCCCAATGACACTTGGGCTCTCAACTCATGGAAGTGAATGTTTACA 699  
Db 603 TTTCTATAAACACACAGTCAACTTGACAGCTGGCCCTCAGGTCAATGGAAGTGAATGTTTTC 662  
QY 700 GAAGACATGGACGGGCATATGTTCCCATCGCACAGTGAAGATGTTGACGTGTAACAG 759  
Db 663 GAAGATACGGCGGGCATACATCCCATCTCGAAGTGAAGATGTTGATGTATAACAG 722  
QY 760 ATCAGATTCTCTGTGTTGTGACTATGTTTCCAGAAAGACGATCGAAATTCATCCACGAA 819  
Db 723 ATCAGATCCCTGTATTCGTGACCAATGTTCCCAAGAAATGACAGGAATGTTGCTGATGAGA 782  
QY 820 CTTTCC-AAAAGATCTCCCAATTAATGTTGATGTTCTGATTCATGATCCTAGCCACTTCC 878  
Db 783 TCTTCTCTCAGAGACCTCCCATCGTCTTCGATGCTCTCAATTCATGATCCAGCCACTTCC 842  
QY 879 TCAATTATTTACCAATTAATGATGAGTGGAGCTTTCGGGGATTAATGACGCTGTTGTTT 938  
Db 843 TCAAGACTCTGCGCATTTTCTTCAAGTGGAACTTTGGGGGACACACTGGCTGTTGTTCT 902  
QY 939 CCACCAATCATCTGTGAATCAACACGATGTTGCTCAATGGAACCTTCAGGCTTAACCTCA 998  
Db 903 CCAACCAATCACTTTGAATCACTTATGCTCAATGGAACCTTCAGGCTTAACCTCA 962  
QY 999 CTGTGAAGCTGCAGACACAGGACCTTG----- 1026  
Db 963 CCGTGCAAACTGCAGTGCCTGGGCGCATGCGCTCCCTTCGCTTCGACTTCGCTCCAC 1022  
QY 1027 -----TCGCGCACCGCCACCCACCGACACCTTCAAAACCCACCCCTT 1070  
Db 1023 CTTCAACTCGCCCTCAGCTCCGCCCTCAGCTTCGCCCACTTATCAACACCTTAGCCCT 1082  
QY 1071 CTTTAGGACCTGCTGGTGAACCCCTCTGGAGCTGATGAGTTCCTGATGAAATGCTCC 1130  
Db 1083 CTTTAATGCTACTGTTTCAATCCATGAGCTGATGACATTTTCAATGAAATGCTCC 1142  
QY 1131 AGATTAAAGATATGGCCCTTTCAAGCCACCATCAATTTAGAGGGAATCTTAGAGG 1190  
Db 1143 GAATAAACAGATATGGCTACTTCAGAGCCACCATCAATTTAGAGGGAATCTTAGAG 1202  
QY 1191 TTAACATCATCAGATGACAGAGCTCTGATGCGGTGCGATGGCTGAAAGCTCCCTAA 1250  
Db 1203 TCAGCATATGAGATAGCAGATGTCCTCATGCCCCACCGCAGCCTGCCAATCTCCTGA 1262

```

; LOCATION: (560)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (564)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (566)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (575)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (599)
; OTHER INFORMATION: Where n is a, c, g or t
;
US-09-123-912-105

Query Match      20.5%; Score 545.6; DB 4; Length 619;
Best Local Similarity 93.2%; Pred. No. 3.2e-108; Indels 1; Gaps 1;
Matches 577; Conservative 0; Mismatches 41;

QY      2010  ACTAGTAGGATAGAAACACTGTGTCGCCAGAGTAGAGAGAGAAGCTACTATTGATTAGAG 2069
DB      2    |||||||
QY      2070  CCTAACCCAGGTTAACTGCGAAGAGAGCGGGGATACCTTTCAGCTTTCATGTAACTGTAT 2129
DB      62    |||||||
QY      2130  GCATAAAGCCAAATGATGCCAGTTCTTAAGATCATGTTCCAAAGCTAACTGAATCCCACTT 2189
DB      122   GCATAAAGCCAAATGATGCCAGTTCTTAAGATCATGTTCCAAAGCTAACTGAATCCCACTT 181
QY      2190  CAATACACACTCATGAACCTCTGTATGGAAACAATAACAGCCCAAGCCCTGTGTGTATGATGT 2249
DB      182   CAATACACACTCATGAACCTCTGTATGGAAACAATAACAGCCCAAGCCCTGTGTGTATGATGT 241
QY      2250  GCACACTTGCAGCTCAGAAAAATACTACTCTCATAAATGGGTGGGAGTATTTTGGTG 2309
DB      242   GCACACTTGCAGCTCANAAAAATACTACTCTCATAAATGGGTGGGAGTATTTTGGTG 301
QY      2310  ACAACCTACTTTGCTTGGCTGAGTGAAGGAATGATTCATATATATTCATTTATTCATGG 2369
DB      302   ACAACCTACTTTGCTTGGCTGAGTGAAGGAATGATTCATATATATTCATTTATTCATGG 361
QY      2370  ACATTTAGTTAGTGTCTTTTATATACAGCGCATGATGCTGAGTGACACTCTTGTGTATAT 2429
DB      362   ACATTTAGTTAGTGTCTTTTATATACAGCGCATGATGCTGAGTGACACTCTTGTGTATAT 421
QY      2430  TTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATCAAAATATTTAAGACTTTCCAAA 2489
DB      422   TTCCAAATTTTGTACAGTCGCTGCACATATTTGAAATC-AATATTTAAGACTTTCCAAA 480
QY      2490  AATTTGGTCCTCGTGTCTTTTCATGCGCACTTGATCAGTAAGGATTTCCCTCTGTTTGGAA 2549
DB      481   AATGAAGTCCTCGTGTCTTTTCATGCGCACTTGATCAGTAAGGATTTCCCTCTGTTTGGTA 540
QY      2550  CTAAAACCATTTACTATATGTTAGACAGACATTTTTTTTTTCTCTCTGAGAAAAAAA 2609
DB      541   CTTAAAAACATCTACTATATGTTNANATGAAATTCCTTTTCCCNCTCCCGAAAAAANA 600
QY      2610  ATGAGGGAAGAGACAAAAA 2628
DB      601   AAGTGTGGGAAAAAAA 619

RESULT 8
US-09-643-597-1105
; Sequence 105, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.

```

1251	Qy	TAGACTTTGTCTGTGACCTGCCAAGGGACATTC	1310
		CGAGGCTGTGTACCACTATTTCTG	
1263	Db	TGGACTTCACTGTGACCTGCAAAAGGGGCCA	1322
		CCCCCATGGAAAGCCTGTACGATCATCTCG	
1311	Qy	ACCCACCTGGAGATACCCAGAAACAGTCTG	1370
		CAGCCCTCTGGGATGTGATGAGATGT	
1323	Db	ACCCACCTGCCAGATCGGCCAGAACCGGTC	1382
		GTGAGCCCTGTGGCTGTGATGGGCTGT	
1371	Qy	GTCTGCTGACTGTGTGAGAGAACTTTCAAT	1430
		GGGTCTGGGACGTACTGTGTGAACCTCACCC	
1383	Db	GCCTGCTGTCTGTGAGAGAGCGTTTCAAT	1442
		GGGTCTGGCACCTACTGTGTGAAATTCACTC	
1431	Qy	TGGGGATGACACAAAGCTGGCTCTCAGAG	1490
		CACCCCTGATTTCTCTCTGACAGAGACC	
1443	Db	TGGGAGATGATCAAGCTGGGCCCTCACAG	1502
		CACCCCTGATCTCTATCTCTCCCTGGCAAGACC	
1491	Qy	CAGCCTCGCCTTTAAGATGGCAACAGTGC	1550
		CCCTGATCTCCGTTGGCTCTTGGCCATAT	
1503	Db	CAGACTCCCTCTGAGAGCAGTGATGGTGT	1562
		CTCTGATCTCCATTTGGCTGGCTGTGTC	
1551	Qy	TTGTCACTGTGATCTCCCTCTTTGGTGTGA	1610
		CAAAAACACAAAGGAATACACCCCAATAGAAA	
1563	Db	TTGTCACTGATGGTTACCATCTTGTCTGTGA	1622
		CAAAAACACAAAGCGTACAAAGCCCAATAGGAA	
1611	Qy	ATAGTCTCGGAATGTGGTCAGAAGCAAA	1670
		AGGCTGAGTGTCTTTCTCAACCGCTGCAAAAG	
1623	Db	ACTGCCCCAGGAACACGGTCAAGGGCA	1682
		AGGCTGAGTGTCTCTCTCAGCCACGCGAAAG	
1671	Qy	CCGTGTTCTTCCCGGAAACACAGAAAGAT	1730
		CCGCTACTCAAAAACCAAGAAATTTAAAG	
1683	Db	CCCCGTCTTCCGAGGAGACCAGAGAAAG	1742
		ATTCATTGTCTCCAGGCAAGACCAAGGACAC	
1731	Qy	GAGTTCTTAAATTTGACCTTGTGTTCTG	1790
		AGCTCACTTTTCAGTGCCTATTGATGTGACA	
1743	Db	TCTAAGTCTTTTGGCCTTCCCTCTGAC	1801
		CAGGAACCCACTCTTCTGTG-CATGTATGTGAGC	
1791	Qy	TGTGCTGGA---GTGCTATTAACTTTTT	1846
		TCCTAAAGATTATTGTTAAATAGATATT	
1802	Db	TGTGCAGAGATGTGGCTGGGAACGTGTG	1861
		TTCTTAAGGAATATTGTTAAATGTATATC	
1847	Qy	GTGCTTTGGGGAAGTTGAATTTTTTAT	1881
		TAGGTTAAA	
1862	Db	GTGCTTAGGAGTGTGTTTAAATAGCAT	1896
		TTTTTAAATAGCATTTTAA	

```

RESULT 7
US-09-123-912-105
; Sequence 105, Application US/09123912A
; Patent No. 6312695
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
; FILE REFERENCE: 210121.455C1
; CURRENT APPLICATION NUMBER: US/09/123,912A
; CURRENT FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: 09/040,802
; PRIOR FILING DATE: 1998-03-18
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 105
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (260)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (527)
; OTHER INFORMATION: Where n is a, c, g or t
; NAME/KEY: modified_base

```

APPLICANT: Bangur, Chaitanya S.  
APPLICANT: Hosken, Nancy  
APPLICANT: Fanger, Gary R.  
APPLICANT: Li, Samuel X.  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Henderson, Robert A.  
APPLICANT: McNeill, Patricia D.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
FILE REFERENCE: 210121.455C11  
CURRENT APPLICATION NUMBER: US/09/643,597  
CURRENT FILING DATE: 2000-08-21  
NUMBER OF SEQ ID NOS: 369  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 105  
LENGTH: 619  
TYPE: DNA  
ORGANISM: Homo sapien  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)...(619)  
OTHER INFORMATION: n = A,T,C or G  
US-09-643-597-105

Query Match 20.5%; Score 545.6; DB 4; Length 619;  
Best Local Similarity 93.2%; Pred. No. 3.2e-108;  
Matches 577; Conservative 0; Mismatches 41; Indels 1; Gaps 1;  
QY 2010 ACTAGTAGGATAGAAACACTGTGTCGCGAGAGTAAAGAGAGAGTAACTATTGATTAGAG 2069  
Db 2 ACTAGTAGGATAGAAACACTGTGTCGCGAGAGTAAAGAGAGAGTAACTATTGATTAGAG 61  
QY 2070 CCTAACCCAGGTTAACTGCAAGAGAGCGGGATCTTTCAGCTTTCATGTAAGTGTAT 2129  
Db 62 CCTAACCCAGGTTAACTGCAAGAGAGCGGGATCTTTCAGCTTTCATGTAAGTGTAT 121  
QY 2130 GCATTAAGCCCAATGTAGTCCAGGTTTCTAAGATCATGTTCAAGCTAACTGAATCCCACTT 2189  
Db 122 GCATTAAGCCCAATGTAGTCCAGGTTTCTAAGATCATGTTCAAGCTAACTGAATCCCACTT 181  
QY 2190 CAATACACTCATGAACCTCTGATGAAACAATTAACAGGCCCAAGCCTGTGGTATGATG 2249  
Db 182 CAATACACTCATGAACCTCTGATGAAACAATTAACAGGCCCAAGCCTGTGGTATGATG 241  
QY 2250 GCACACTTGTAGACTCAGAAAAATACTACTCTCATAAATGGTGGAGTATTTGGTG 2309  
Db 242 GCACACTTGTAGACTCAGAAAAATACTACTCTCATAAATGGTGGAGTATTTGGTG 301  
QY 2310 ACAACCTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCATTTATTCATGG 2369  
Db 302 ACAACCTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCATTTATTCATGG 361  
QY 2370 ACATTTAGTTAGTGTCTTTTATATACAGGCATGCTGAGTGACACTCTTGTGTATAT 2429  
Db 362 ACATTTAGTTAGTGTCTTTTATATACAGGCATGCTGAGTGACACTCTTGTGTATAT 421  
QY 2430 TTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATCAAATATTAAGACTTCCAAA 2489  
Db 422 TTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATCAAATATTAAGACTTCCAAA 480  
QY 2490 AATTTGGTCCCTGGTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTTGGAA 2549  
Db 481 AATGAAGTCCCTGGTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTTGGAA 540  
QY 2550 CTAAACCAATTTACTATATGTTAGACAAGACATTTTTTTTTTTCCTCTCTGAAAAAAA 2609  
Db 541 CTAAACCAATTTACTATATGTTAGACAAGACATTTTTTTTTTTCCTCTCTGAAAAAAA 600  
QY 2610 ATGAGGGAAGAGACAAAA 2628  
Db 601 AAGTGTGGGAAAAAAA 619

RESULT 9  
US-09-643-597-121  
Sequence 121, Application US/09643597  
Patent No. 6426072  
GENERAL INFORMATION:  
APPLICANT: Wang, Tongtong  
APPLICANT: Fan, Liqun  
APPLICANT: Kalos, Michael D.  
APPLICANT: Bangur, Chaitanya S.  
APPLICANT: Hosken, Nancy  
APPLICANT: Fanger, Gary R.  
APPLICANT: Li, Samuel X.  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Henderson, Robert A.  
APPLICANT: McNeill, Patricia D.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
FILE REFERENCE: 210121.455C11  
CURRENT APPLICATION NUMBER: US/09/643,597  
CURRENT FILING DATE: 2000-08-21  
NUMBER OF SEQ ID NOS: 369  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 121  
LENGTH: 619  
TYPE: DNA  
ORGANISM: Homo sapien  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)...(619)  
OTHER INFORMATION: n = A,T,C or G  
US-09-643-597-121

Query Match 20.5%; Score 545.6; DB 4; Length 619;  
Best Local Similarity 93.2%; Pred. No. 3.2e-108;  
Matches 577; Conservative 0; Mismatches 41; Indels 1; Gaps 1;  
QY 2010 ACTAGTAGGATAGAAACACTGTGTCGCGAGAGTAAAGAGAGAGTAACTATTGATTAGAG 2069  
Db 2 ACTAGTAGGATAGAAACACTGTGTCGCGAGAGTAAAGAGAGAGTAACTATTGATTAGAG 61  
QY 2070 CCTAACCCAGGTTAACTGCAAGAGAGCGGGATCTTTCAGCTTTCATGTAAGTGTAT 2129  
Db 62 CCTAACCCAGGTTAACTGCAAGAGAGCGGGATCTTTCAGCTTTCATGTAAGTGTAT 121  
QY 2130 GCATTAAGCCCAATGTAGTCCAGGTTTCTAAGATCATGTTCAAGCTAACTGAATCCCACTT 2189  
Db 122 GCATTAAGCCCAATGTAGTCCAGGTTTCTAAGATCATGTTCAAGCTAACTGAATCCCACTT 181  
QY 2190 CAATACACTCATGAACCTCTGATGAAACAATTAACAGGCCCAAGCCTGTGGTATGATG 2249  
Db 182 CAATACACTCATGAACCTCTGATGAAACAATTAACAGGCCCAAGCCTGTGGTATGATG 241  
QY 2250 GCACACTTGTAGACTCAGAAAAATACTACTCTCATAAATGGTGGAGTATTTGGTG 2309  
Db 242 GCACACTTGTAGACTCAGAAAAATACTACTCTCATAAATGGTGGAGTATTTGGTG 301  
QY 2310 ACAACCTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCATTTATTCATGG 2369  
Db 302 ACAACCTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCATTTATTCATGG 361  
QY 2370 ACATTTAGTTAGTGTCTTTTATATACAGGCATGCTGAGTGACACTCTTGTGTATAT 2429  
Db 362 ACATTTAGTTAGTGTCTTTTATATACAGGCATGCTGAGTGACACTCTTGTGTATAT 421  
QY 2430 TTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATCAAATATTAAGACTTCCAAA 2489  
Db 422 TTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATCAAATATTAAGACTTCCAAA 480  
QY 2490 AATTTGGTCCCTGGTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTTGGAA 2549  
Db 481 AATGAAGTCCCTGGTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTTGGTA 540

```
QY 2550 CTAAGGCACTTACTATATGTTAGACAGACATTTTTTTTTTCCCTTCCCTGGAAGAAA 2609
Db 541 CTTAAACATCTACTATATGTTNANATGAAATTCCTTTCCCNCCCTCCGGAAGAAAANA 600
QY 2610 ATGAGGGAAGAGACAAAA 2628
Db 601 AAGTGGTGGGGAAGAAA 619

RESULT 10
US-09-480-884A-105
; Sequence 105, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 105
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(619)
; OTHER INFORMATION: n = A,T,C or G
US-09-480-884A-105

Query Match 20.5%; Score 545.6; DB 4; Length 619;
Best Local Similarity 93.2%; Pred. No. 3.2e-108;
Matches 577; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 2010 ACTAGTAGATAGAAACACTGTGTCAGAGAGTAGAGAGAGAGTACTATTGATTAGAG 2069
Db 2 ACTAGTAGATAGAAACACTGTGTCAGAGAGTAGAGAGAGTACTATTGATTAGAG 61
QY 2070 CCTAACCCAGGTTAACTGCAAGAGAGGCGGGATCTTTTCAGCTTTCCATGTAACGTAT 2129
Db 62 CCTAACCCAGGTTAACTGCAAGAGAGGCGGGATCTTTTCAGCTTTCCATGTAACGTAT 121
QY 2130 GCATAAGCCATGTAGTCCAGTTTCTAAGATCATGTTTCCAAAGCTAACTGAATCCCACIT 2189
Db 122 GCATAAGCCATGTAGTCCAGTTTCTAAGATCATGTTTCCAAAGCTAACTGAATCCCACIT 181
QY 2190 CAATACACACTCATGAACCTCTGATGGAACAATAACAGGCCCAAGCCTGTGTATGATGT 2249
Db 182 CAATACACACTCATGAACCTCTGATGGAACAATAACAGGCCCAAGCCTGTGTATGATGT 241
QY 2250 GCACACTTGTAGACTCAGAAAAAATACTACTCTCATAAATGGGTGGGAGTATTTGGTG 2309
Db 242 GCACACTTGTAGACTCAGAAAAAATACTACTCTCATAAATGGGTGGGAGTATTTGGTG 301
QY 2310 ACNACCTACTTCTGGCTGAGTGAAGGAATGATATTCATATTTCAITTTATTCATGG 2369
Db 302 ACNACCTACTTCTGGCTGAGTGAAGGAATGATATTCATATTTCAITTTATTCATGG 361
QY 2370 ACATTTAGTTAGTGTCTTTTATATACCAGGCATGCTGAGTGACACTCTTGTGTATAT 2429
Db 362 ACATTTAGTTAGTGTCTTTTATATACCAGGCATGCTGAGTGACACTCTTGTGTATAT 421
QY 2430 TTCCAAATTTTGTAGTCGCTGCACATATTTGAAATCAAATATTAAGACITTCACAA 2489
Db 422 TTCCAAATTTTGTAGTCGCTGCACATATTTGAAATCAAATATTAAGACITTCACAA 480

QY 2490 AATTGGTCCCTGGTTTTTTCATGCCAACTTGATCAGTAAGATTTCCCTCTCTGTTGAA 2549
Db 481 AATGAAGTCCCTGGTTTTTTCATGCCAACTTGATCAGTAAGATTTCCCTCTCTGTTGTA 540
QY 2550 CTAAGGCACTTACTATATGTTAGACAGACATTTTTTTTTTCCCTTCCCTGGAAGAAA 2609
Db 541 CTTAAACATCTACTATATGTTNANATGAAATTCCTTTCCCNCCCTCCGGAAGAAAANA 600
QY 2610 ATGAGGGAAGAGACAAAA 2628
Db 601 AAGTGGTGGGGAAGAAA 619

RESULT 11
US-09-480-884A-121
; Sequence 121, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 121
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(619)
; OTHER INFORMATION: n = A,T,C or G
US-09-480-884A-121

Query Match 20.5%; Score 545.6; DB 4; Length 619;
Best Local Similarity 93.2%; Pred. No. 3.2e-108;
Matches 577; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 2010 ACTAGTAGATAGAAACACTGTGTCAGAGAGTAGAGAGAGTACTATTGATTAGAG 2069
Db 2 ACTAGTAGATAGAAACACTGTGTCAGAGAGTAGAGAGAGTACTATTGATTAGAG 61
QY 2070 CCTAACCCAGGTTAACTGCAAGAGAGGCGGGATCTTTTCAGCTTTCCATGTAACGTAT 2129
Db 62 CCTAACCCAGGTTAACTGCAAGAGAGGCGGGATCTTTTCAGCTTTCCATGTAACGTAT 121
QY 2130 GCATAAGCCATGTAGTCCAGTTTCTAAGATCATGTTTCCAAAGCTAACTGAATCCCACIT 2189
Db 122 GCATAAGCCATGTAGTCCAGTTTCTAAGATCATGTTTCCAAAGCTAACTGAATCCCACIT 181
QY 2190 CAATACACACTCATGAACCTCTGATGGAACAATAACAGGCCCAAGCCTGTGTATGATGT 2249
Db 182 CAATACACACTCATGAACCTCTGATGGAACAATAACAGGCCCAAGCCTGTGTATGATGT 241
QY 2250 GCACACTTGTAGACTCAGAAAAAATACTACTCTCATAAATGGGTGGGAGTATTTGGTG 2309
Db 242 GCACACTTGTAGACTCAGAAAAAATACTACTCTCATAAATGGGTGGGAGTATTTGGTG 301
QY 2310 ACNACCTACTTCTGGCTGAGTGAAGGAATGATATTCATATTTCAITTTATTCATGG 2369
Db 302 ACNACCTACTTCTGGCTGAGTGAAGGAATGATATTCATATTTCAITTTATTCATGG 361
QY 2370 ACATTTAGTTAGTGTCTTTTATATACCAGGCATGCTGAGTGACACTCTTGTGTATAT 2429
Db 362 ACATTTAGTTAGTGTCTTTTATATACCAGGCATGCTGAGTGACACTCTTGTGTATAT 421
QY 2430 TTCCAAATTTTGTAGTCGCTGCACATATTTGAAATCAAATATTAAGACITTCACAA 2489
Db 422 TTCCAAATTTTGTAGTCGCTGCACATATTTGAAATCAAATATTAAGACITTCACAA 2489
```

Db 422 TTCCAAATTTTGTACAGTGGCTGCACATATTTGAAATC-ATATATTAAGACTTCCAAA 480  
QY 2490 AATTGGTCCCTCGTTTTCATGGCACTTGATCAGTAAGGATTTCCCTCTCTGTTGGAA 2549  
Db 481 AATGAAGTCCCTCGTTTTCATGGCACTTGATCAGTAAGGATTTCCCTCTCTGTTGGTA 540  
QY 2550 CTAAACCAATTTACTATATGATAGCAAGACATTTTTCCTTCTGCTGAAAAAAA 2609  
Db 541 CTAAACCACTACTATATGTTNANATGAAATTCCTTTCCCNCTCCCGAAAAANA 600  
QY 2610 ATGAGGGAAGAGACAAAA 2628  
Db 601 AAGTGGTGGGAAAAAAA 619

## RESULT 12

US-09-542-615A-105

; Sequence 105, Application US/09542615A

; Patent No. 6518256

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosken, Nancy A.

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY

; FILE REFERENCE: 210121.455C8

; CURRENT APPLICATION NUMBER: US/09/542,615A

; NUMBER OF SEQ ID NOS: 350

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 105

; LENGTH: 619

; TYPE: DNA

; ORGANISM: Homo sapien

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1) - (619)

; OTHER INFORMATION: n = A,T,C or G

US-09-542-615A-105

Query Match 20.5%; Score 545.6; DB 4; Length 619;

Best Local Similarity 93.2%; Pred. No. 3.2e-108;

Matches 577; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 2010 ACTAGTAGGATAGAAACACTGTGTCGCGAGAGTAAGGAGAGAGCTACTATTGATTAGAG 2069  
Db 2 ACTAGTAGGATAGAAACACTGTGTCGCGAGAGTAAGGAGAGAGCTACTATTGATTAGAG 61

QY 2070 CCTAACCCAGGTTAACTGTCAGAGAGAGCGGGGATCTTTTCAGCTTTCATGTAAGTAT 2129  
Db 62 CCTAACCCAGGTTAACTGTCAGAGAGAGCGGGGATCTTTTCAGCTTTCATGTAAGTAT 121

QY 2130 GCATAAGCCAAATGATGTCAGATTTCTAAGATCATGTTCCAAAGTAACTGATCCACTT 2189  
Db 122 GCATAAGCCAAATGATGTCAGATTTCTAAGATCATGTTCCAAAGTAACTGATCCACTT 181

QY 2190 CATACACACTCATGAACTCCTGATGGAACAATAACAGGCCCAAGCCTGTGGTATGATGT 2249  
Db 182 CATACACACTCATGAACTCCTGATGGAACAATAACAGGCCCAAGCCTGTGGTATGATGT 241

QY 2250 GCACACTTGTAGACTCAGAAAAATACTACTCTATAAATGGTGGGAGTATTTGGTG 2309  
Db 242 GCACACTTGTAGACTCAGAAAAATACTACTCTATAAATGGTGGGAGTATTTGGTG 301

QY 2310 ACAACCTACTTGTGCTGGCTGAGTGAAGGAATGATATTCATATATTCATTTCCATGG 2369  
Db 302 ACAACCTACTTGTGCTGGCTGAGTGAAGGAATGATATTCATATATTCATTTCCATGG 361

QY 2370 ACATTTAGTTAGTGTCTTTTATATACCAGGCATGATCTGAGTGACACTCTTGTGTATAT 2429

Db 362 ACATTTAGTTAGTGTCTTTTATATACCAGGCATGATCTGAGTGACACTCTTGTGTATAT 421  
QY 2430 TTCCAAATTTTCTATAGTGGCTGCACATATTTGAAATC-ATATATTAAGACTTCCAAA 2489  
Db 422 TTCCAAATTTTGTACAGTGGCTGCACATATTTGAAATC-ATATATTAAGACTTCCAAA 480  
QY 2490 AATTGGTCCCTCGTTTTCATGGCACTTGATCAGTAAGGATTTCCCTCTCTGTTGGAA 2549  
Db 481 AATGAAGTCCCTCGTTTTCATGGCACTTGATCAGTAAGGATTTCCCTCTCTGTTGGTA 540  
QY 2550 CTAAACCAATTTACTATATGTTAGCAAGACATTTTTCCTTCTGCTGAAAAAAA 2609  
Db 541 CTAAACCACTACTATATGTTNANATGAAATTCCTTTCCCNCTCCCGAAAAANA 600  
QY 2610 ATGAGGGAAGAGACAAAA 2628  
Db 601 AAGTGGTGGGAAAAAAA 619

## RESULT 13

US-09-542-615A-121

; Sequence 121, Application US/09542615A

; Patent No. 6518256

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hosken, Nancy A.

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY

; FILE REFERENCE: 210121.455C8

; CURRENT APPLICATION NUMBER: US/09/542,615A

; CURRENT FILING DATE: 2000-04-14

; NUMBER OF SEQ ID NOS: 350

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 121

; LENGTH: 619

; TYPE: DNA

; ORGANISM: Homo sapien

; FEATURE:

; NAME/KEY: misc feature

; LOCATION: (1) - (619)

; OTHER INFORMATION: n = A,T,C or G

US-09-542-615A-121

Query Match 20.5%; Score 545.6; DB 4; Length 619;

Best Local Similarity 93.2%; Pred. No. 3.2e-108;

Matches 577; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 2010 ACTAGTAGGATAGAAACACTGTGTCGCGAGAGTAAGGAGAGAGCTACTATTGATTAGAG 2069  
Db 2 ACTAGTAGGATAGAAACACTGTGTCGCGAGAGTAAGGAGAGAGCTACTATTGATTAGAG 61

QY 2070 CCTAACCCAGGTTAACTGTCAGAGAGAGCGGGGATCTTTTCAGCTTTCATGTAAGTAT 2129  
Db 62 CCTAACCCAGGTTAACTGTCAGAGAGAGCGGGGATCTTTTCAGCTTTCATGTAAGTAT 121

QY 2130 GCATAAGCCAAATGATGTCAGATTTCTAAGATCATGTTCCAAAGTAACTGATCCCACTT 2189  
Db 122 GCATAAGCCAAATGATGTCAGATTTCTAAGATCATGTTCCAAAGTAACTGATCCCACTT 181

QY 2190 CATACACACTCATGAACTCCTGATGGAACAATAACAGGCCCAAGCCTGTGGTATGATGT 2249  
Db 182 CATACACACTCATGAACTCCTGATGGAACAATAACAGGCCCAAGCCTGTGGTATGATGT 241

QY 2250 GCACACTTGTAGACTCAGAAAAATACTACTCTATAAATGGTGGGAGTATTTGGTG 2309  
Db 242 GCACACTTGTAGACTCAGAAAAATACTACTCTATAAATGGTGGGAGTATTTGGTG 301

QY 2310 ACAACCTACTTGTGCTGGCTGAGTGAAGGAATGATATTCATATATTCATTTCCATGG 2369

Db 182 CAATACACACTCATGAACTCTGTGATGGAAACAATTAACAGGCCCAAGCCCTGTGTATGATGT 241  
QY 2250 GCACACTTGTAGACTCAGAAAAAATACTACTCTCATAAAATGGGTGGAGTATTTTGGTG 2309  
Db 242 GCACACTTGTAGACTCANAAAAAATACTACTCTCATAAAATGGGTGGAGTATTTTGGTG 301  
QY 2310 ACAACCTACTTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2369  
Db 302 ACAACCTACTTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 361  
QY 2370 ACATTTAGTGTAGCTTTTATATATACACAGGCATGATGCTGAGTGACACTCTTGTGTATAT 2429  
Db 362 ACATTTAGTGTAGCTTTTATATATACACAGGCATGATGCTGAGTGACACTCTTGTGTATAT 421  
QY 2430 TTCCAAATTTTGTATAGTCTGTCACATATTTTGAATCAAAATTAAGACTTTCCAAA 2489  
Db 422 TTCCAAATTTTGTATAGTCTGTCACATATTTTGAATCAAAATTAAGACTTTCCAAA 480  
QY 2490 AATTTGGTCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2549  
Db 481 AATGAAGTCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 540  
QY 2550 CTAACACCATTTACTATATGTTAGCAAGACATTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2609  
Db 541 CTTAAACATCTACTATATGTTTANATGAAATTCCTTTTCCCNCTCCCGAAAAANA 600  
QY 2610 ATGAGGGAAGACAAAAA 2628  
Db 601 AAGTGGTGGGAAAAA 619  
RESULT 15  
US-09-606-421B-121  
; Sequence 121, Application US/09606421B  
; Patent No. 6531315  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C9  
; CURRENT FILING DATE: 2000-06-28  
; NUMBER OF SEQ ID NOS: 358  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 121  
; LENGTH: 619  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(619)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-606-421B-121  
Query Match 20.5%; Score 545.6; DB 4; Length 619;  
Best Local Similarity 93.2%; Pred. No. 3.2e-108;  
Matches 577; Conservative 0; Mismatches 41; Indels 1; Gaps 1;  
QY 2010 ACTAGTAGGATAGAAACACTGTGTCCGAGAGTAAGGAGAGAGCTACTACTATTGATTAGAG 2069  
Db 2 ACTAGTAGGATAGAAACACTGTGTCCGAGAGTAAGGAGAGAGCTACTACTATTGATTAGAG 61  
QY 2070 CTTAACCCAGGTTAACTGCAAGAGAGCGGGATCTTTTCCAGCTTTCATGTACTGTAT 2129  
Db 62 CTTAACCCAGGTTAACTGCAAGAGAGCGGGATCTTTTCCAGCTTTCATGTACTGTAT 121  
QY 2130 GCATAAGCCAACTGCTCAGTTTCTAAGATCATGTTTCAAGCTAACTGAATCCCACTT 2189  
Db 122 GCATAAGCCAACTGCTCAGTTTCTAAGATCATGTTTCAAGCTAACTGAATCCCACTT 181  
QY 2190 CAATACACACTCATGAATCTCTGTATGGAACAATTAACAGGCCCAAGCCCTGTGTATGATGT 2249  
Db 62 CTTAACCCAGGTTAACTGCAAGAGAGCGGGATCTTTTCCAGCTTTCATGTACTGTAT 121

Db 302 ACAACCTACTTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 361  
QY 2370 ACATTTAGTGTAGCTTTTATATATACACAGGCATGATGCTGAGTGACACTCTTGTGTATAT 2429  
Db 362 ACATTTAGTGTAGCTTTTATATATACACAGGCATGATGCTGAGTGACACTCTTGTGTATAT 421  
QY 2430 TTCCAAATTTTGTATAGTCTGTCACATATTTTGAATCAAAATTAAGACTTTCCAAA 2489  
Db 422 TTCCAAATTTTGTATAGTCTGTCACATATTTTGAATCAAAATTAAGACTTTCCAAA 480  
QY 2490 AATTTGGTCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2549  
Db 481 AATGAAGTCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 540  
QY 2550 CTAACACCATTTACTATATGTTAGCAAGACATTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2609  
Db 541 CTTAAACATCTACTATATGTTTANATGAAATTCCTTTTCCCNCTCCCGAAAAANA 600  
QY 2610 ATGAGGGAAGACAAAAA 2628  
Db 601 AAGTGGTGGGAAAAA 619  
RESULT 14  
US-09-606-421B-105  
; Sequence 105, Application US/09606421B  
; Patent No. 6531315  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.455C9  
; CURRENT FILING DATE: 2000-06-28  
; NUMBER OF SEQ ID NOS: 358  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 105  
; LENGTH: 619  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(619)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-606-421B-105  
Query Match 20.5%; Score 545.6; DB 4; Length 619;  
Best Local Similarity 93.2%; Pred. No. 3.2e-108;  
Matches 577; Conservative 0; Mismatches 41; Indels 1; Gaps 1;  
QY 2010 ACTAGTAGGATAGAAACACTGTGTCCGAGAGTAAGGAGAGAGCTACTACTATTGATTAGAG 2069  
Db 2 ACTAGTAGGATAGAAACACTGTGTCCGAGAGTAAGGAGAGAGCTACTACTATTGATTAGAG 61  
QY 2070 CTTAACCCAGGTTAACTGCAAGAGAGCGGGATCTTTTCCAGCTTTCATGTACTGTAT 2129  
Db 62 CTTAACCCAGGTTAACTGCAAGAGAGCGGGATCTTTTCCAGCTTTCATGTACTGTAT 121  
QY 2130 GCATAAGCCAACTGCTCAGTTTCTAAGATCATGTTTCAAGCTAACTGAATCCCACTT 2189  
Db 122 GCATAAGCCAACTGCTCAGTTTCTAAGATCATGTTTCAAGCTAACTGAATCCCACTT 181  
QY 2190 CAATACACACTCATGAATCTCTGTATGGAACAATTAACAGGCCCAAGCCCTGTGTATGATGT 2249  
Db 62 CTTAACCCAGGTTAACTGCAAGAGAGCGGGATCTTTTCCAGCTTTCATGTACTGTAT 121



```
QY 2130 GCATAAGCCAAATGTAAGTCCAGTTTCTAAGATCATGTTCCAGCTAACTGAATCCCACTT 2189
Db |||
QY 122 GCATAAAGCCAAATGTAAGTCCAGTTTCTAAGATCATGTTCCAGCTAACTGAATCCCACTT 181
Db |||
QY 2190 CAATACACACTCATGTAAGTCCCTGATGGAACATATAACAGGCCCAAGCCTGTGTATGATGT 2249
Db |||
QY 182 CAATACACACTCATGTAAGTCCCTGATGGAACATATAACAGGCCCAAGCCTGTGTATGATGT 241
Db |||
QY 2250 GCACACTTGTGTAAGTCCAGAAAAATACTACTCTCATTAATGGTGGAGTATTTGGTG 2309
Db |||
QY 242 GCACACTTGTGTAAGTCCAGAAAAATACTACTCTCATTAATGGTGGAGTATTTGGTG 301
Db |||
QY 2310 ACAACCTACTTTGCTGGTGGTGAAGGAATGATTCATATATTTCAATTTTCCATGG 2369
Db |||
QY 302 ACAACCTACTTTGCTGGTGGTGAAGGAATGATTCATATATTTCAATTTTCCATGG 361
Db |||
QY 2370 ACATTTAGTTAGTGTCTTTTATATACAGGCATGATGCTGAGTGACACTCTTGTGTATAT 2429
Db |||
QY 362 ACATTTAGTTAGTGTCTTTTATATACAGGCATGATGCTGAGTGACACTCTTGTGTATAT 421
Db |||
QY 2430 TTCCAAATTTTGTATAGTCGCTGCACATATTTGAATCARAAATTAAGACTTTCCAAA 2489
Db |||
QY 422 TTCCAAATTTTGTACAGTCGCTGCACATATTTGAATC-ATATATTAAGACTTTCCAAA 480
Db |||
QY 2490 AATTGGTCCCTGGTTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTTTGGAA 2549
Db |||
QY 481 AATGAAGTCCCTGGTTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTTTGGTA 540
Db |||
QY 2550 CTAAACCACTTACTATATGTTAGACAAGACATTTTTTTTTTCTTCTTCTGAAAAAAA 2609
Db |||
QY 541 CTAAACCACTTACTATATGTTAGACAAGACATTTTTTTTTTCTTCTTCTGAAAAAAA 600
Db |||
QY 2610 ATGAGGGAAGAGACAAAAA 2628
Db |||
QY 601 AAGTGGTGGGAAAAAAA 619
Db |||
```

Search completed: September 10, 2004, 10:01:25  
Job time : 218 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: September 10, 2004, 09:57:48 ; Search time 200 Seconds  
(without alignments)  
7518.589 Million cell updates/sec

Title: US-10-039-272-1  
Perfect score: 4850  
Sequence: 1 cggcagaggccagagga.....aaaaaaaaaaaaaaaaaaaaa 2661

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1586107 segs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool\_p/US10039272/runat\_07092004\_144222\_21609/app\_query.fasta\_1.2823  
-DB=A\_Geneseq\_29Jan04 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS-human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptp -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10039272.cgn\_1\_1\_265@runat\_07092004\_144222\_21609 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A\_Geneseq\_29Jan04:  
1: Geneseq1980s:  
2: Geneseq1990s:  
3: Geneseq2000s:  
4: Geneseq2001s:  
5: Geneseq2002s:  
6: Geneseq2003s:  
7: Geneseq2003bs:  
8: Geneseq2004s:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2999	61.8	560	2 AAW35382	Aaw35382 Murine me
2	2999	61.8	560	3 AAB11329	Aab11329 Human lun
3	2999	61.8	560	5 ABB78200	Abb78200 Amino aci
4	2999	61.8	560	5 ABB74961	Abb74961 Human lun
5	2999	61.8	560	5 ABP61881	Abp61881 Human lun
6	2999	61.8	560	6 ABUS6592	Abus6592 Lung canc
7	2999	61.8	560	6 ABG72962	Abg72962 Human ost
8	2999	61.8	560	6 ABU70852	Abu70852 Human adi
9	2999	61.8	560	7 ADA28315	Ada28315 Human lun
10	2980	61.4	572	7 ADD78235	Add78235 Human CGD

11	2979	61.4	572	5	AAU83612	Human PRO
12	2979	61.4	572	6	ABU80759	Human PRO
13	2979	61.4	572	6	ABO33725	Novel hum
14	2979	61.4	572	6	ABU82068	Novel hum
15	2979	61.4	572	6	ABJ72248	Human PRO
16	2979	61.4	572	6	ABJ72376	Human PRO
17	2979	61.4	572	6	ABO34271	Human sec
18	2979	61.4	572	7	ABJ72078	Human mem
19	2979	61.4	572	7	ADB83532	Novel hum
20	2979	61.4	572	7	ADB80638	Novel hum
21	2979	61.4	572	7	ADB73179	Novel hum
22	2979	61.4	572	7	ADB78261	Novel hum
23	2979	61.4	572	7	ADB84909	Human PRO
24	2979	61.4	572	7	ADB78015	Novel hum
25	2979	61.4	572	7	ADB87081	Human PRO
26	2979	61.4	572	7	ADB84663	Human PRO
27	2979	61.4	572	7	ADB83778	Novel hum
28	2979	61.4	572	7	ADB72933	Novel hum
29	2979	61.4	572	7	ADC36771	Human PRO
30	2979	61.4	572	7	ADC21761	Human PRO
31	2979	61.4	572	7	ADC49792	Novel hum
32	2979	61.4	572	7	ADC48991	Novel hum
33	2979	61.4	572	7	ADC49508	Novel hum
34	2979	61.4	572	7	ADC47369	Novel hum
35	2979	61.4	572	7	ADC47114	Novel hum
36	2979	61.4	572	7	ADC77989	Novel hum
37	2979	61.4	572	7	ADD06224	Novel hum
38	2979	61.4	572	7	ADC77743	Novel hum
39	2979	61.4	572	7	ADD50706	Novel hum
40	2979	61.4	572	7	ADD50952	Novel hum
41	2979	61.4	572	7	ADD50433	Human PRO
42	2979	61.4	572	7	ADD50187	Human PRO
43	2979	61.4	572	7	ADD51198	Novel hum
44	2979	61.4	572	8	ADC48745	Novel hum
45	2979	61.4	572	8	ADE20916	Novel hum

## ALIGNMENTS

RESULT 1

AAW35382

ID AAW35382 standard; protein; 560 AA.

AC AAW35382;

XX XX

DT 26-FEB-1998 (first entry)

XX XX

DE Murine metastatic nucleic acid sequence product.

XX XX

KW Mouse; murine; tumour; cancer; metastatic sequence; detection; diagnosis; treatment; metastasis; hyperplasia; dysplasia; hypertrophy; screening.

XX XX

OS Mus musculus.

XX XX

PN WO9718454-A2.

XX XX

PD 22-MAY-1997.

XX XX

PF 15-NOV-1996; 96WO-US018567.

XX XX

PR 16-NOV-1995; 95US-0006838P.

PR 30-JAN-1996; 96US-00594031.

XX XX

(THOM/) THOMPSON T.

PA PA

XX XX

PI Thompson T;

XX XX

DR WPI; 1997-289397/26.

XX XX

Identifying tumour metastatic sequences - by introducing transfected cells into host mammal and analysing primary and metastatic sequences by differential display PCR.

PT PT

XX XX



DT 21-FEB-2001 (first entry)  
 XX Human lung cancer-associated protein L528S.  
 DE Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;  
 XX vaccine; detection.  
 KW Homo sapiens.  
 XX W0200061612-A2.  
 XX 19-OCT-2000.  
 XX 03-APR-2000; 2000WO-US008896.  
 XX 02-APR-1999; 99US-00285479.  
 XX 17-DEC-1999; 99US-00466396.  
 XX 30-DEC-1999; 99US-00476496.  
 XX 10-JAN-2000; 2000US-00480884.  
 XX 22-FEB-2000; 2000US-00510376.  
 XX (CORI-) CORIXA CORP.  
 XX Wang T, Fan L;  
 XX WPI; 2000-628399/60.  
 XX Isolated polypeptide comprising an immunogenic portion of a lung tumor  
 PT protein is used for detecting and monitoring progression of lung cancer  
 PT in a patient.  
 XX Claim 3; Page 203-204; 261pp; English.  
 XX This invention describes a novel isolated polypeptide (I) which  
 CC comprising an immunogenic portion of a lung tumor protein or variant (P2)  
 CC which have cytostatic activity. The polypeptides and polynucleotides are  
 CC used in compositions and vaccines to inhibit the development of cancer,  
 CC especially lung cancer, in a patient. Methods described in the invention  
 CC can be used to monitor the progression of a cancer by carrying out the  
 CC detection at subsequent time points and comparing the results from the  
 CC different time points. CD4+ and/or CD8+ T-Cells isolated from a patient  
 CC are treated with P2, polynucleotides encoding P2 or antigen presenting  
 CC cells expressing P2 and then administered to the patient to inhibit  
 CC development of cancer  
 XX Sequence 560 AA;  
 SQ

Alignment Scores:  
 Pred. No.: 4.82e-295 Length: 560  
 Score: 2999.00 Matches: 559  
 Percent Similarity: 99.82% Conservative: 0  
 Best Local Similarity: 99.82% Mismatches: 1  
 Query Match: 61.84% Indels: 1  
 DB: 3 Gaps: 0

US-10-039-272-1 (1-2661) x AAB11329 (1-560)  
 QY 60 ATGAATGCTCTACTATTTCCTGGATTCTGCTCTGCTGCTGCAAGATTGCCACTTGAT 119  
 DB 1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuLeuAlaAlaArgLeuProLeuAsp 20  
 QY 120 GCCGCCAAACGANTTCATGATGCTCTGGCAATGAAGACCTTCTGCTTACATGAGGAG 179  
 DB 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40  
 QY 180 CACAATCAATTAATGCTGGTCTTCTGATGAAATGACTGGAAATGAAATCACTACCCA 239  
 DB 41 HisAsnGlnLeuAsnGlyTyrSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60  
 QY 240 GTGTGGAAGCGGGAGACATGAGTGGAAAAAATCTCTGGGAAGGAGCGCTGTGCAGGCG 299  
 DB 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla 80

300 GTCTGACCACTGACTCACCAGCCCTCTGTGGGCTCAAATATAACATTTGCGGTGAACCTG 359  
 DB 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100  
 QY 360 ATATTTCCTAGATGCCAAAAGAGATGCCAATGCCAATAGCATAGTCTATGAGAGAACTGC 419  
 DB 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120  
 QY 420 AGAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAACTGGACAGCATGTGCAGAG 479  
 DB 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140  
 QY 480 GACAGTACGGGGAAAATGGCACCGCCAAAGCCATCATACGCTTCCCTGATGGGAAA 539  
 DB 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisHisAsnValPheProAspGlyLys 160  
 QY 540 CTTTTTCTCACCACCCCGGATGAGAGATGGAAATTCATCTACGCTCTCCACACACTT 599  
 DB 161 ProPheProHisHisProGlyTyrArgArgTrpAsnPheIleTyrValPheHisThrIleu 180  
 QY 600 GGTCAGTATTTCAGAAAATTGGGACGATGTTTCACTGAGAGTTCCTGTGAACACACGCCAAT 659  
 DB 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200  
 QY 660 GTGACACTTGGGCTCAACTCATGGAAGTACTGCTTACAGAAACATGAGACGGCATAT 719  
 DB 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220  
 QY 720 GTTCCCATCCGACAAAGTGAAGATGCTGCTGTGTAAACAGATCAGATCTCTGCTTTGTG 779  
 DB 221 ValProIleAlaGlnValLysAspValTyrValThrAspGlnIleProValPheVal 240  
 QY 780 ACTATGTTCCAGAAAGACGATCGAAATTCATCCGACGAAACCTTCCG-AAAGATCTCCCC 838  
 DB 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260  
 QY 839 ATTATGTTGATGCTCTGATTCATGATCTTACGACCTTCCCTCAATTAATTTACCATTAAC 898  
 DB 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280  
 QY 899 TACAAGTGGAGCTCGGGGATATACGCGCTGTTGTTTCCACCATCATCTACTGTGAAT 958  
 DB 281 TyrIleTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300  
 QY 959 CACACGTATGCTCAATGGAACCTTCAGCCTTAACCTCACTCTGTGAAAGCTGCAGCACCA 1018  
 DB 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320  
 QY 1019 GGACCTTTCGGCACCGCCACCCACACCCAGACCTTCAAAACCCACCTCTTTTAGGA 1078  
 DB 321 GlyProCysProProProProProProProProProProProProProProProSerLeuGly 340  
 QY 1079 CTTGCTGGTGACAAACCCCTGGAGCTGAGTAGATTCTCTGATGAAACCTGCCAGATTAAAC 1138  
 DB 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360  
 QY 1139 AGATATGGCCACTTCAAGCCACCATCACAATTTGTAGAGGAATCTTAGAGGTTAAATC 1198  
 DB 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380  
 QY 1199 ATCCAGATGACAGCGCTCTGATGCGGCTGCGATGGCCTGCAAGCTCCCTTAATAGACTTT 1258  
 DB 381 IleGlnMetThrAspValLeuMetProValProValProValProValProValProValPhe 400  
 QY 1259 GTCGTGACCTGCCAAGGAGCATTCACCGAGGTCTGTACCATCATCTTCTGTGACCCACC 1318  
 DB 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420  
 QY 1319 TGCAGATCACCAGAACACAGTCTCTGAGCCCTGTGATGTGTGATGATGATGTGTCTGCTG 1378  
 DB 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440  
 QY 1379 ACTGTGAGACGAACCTTCAATGGGTCTGGACGTACTGTGTGAACCTCACCCCTGGGGAT 1438

```

Db 441 ThrValArgAGThrPheAsnGlySerGlyThrTyrcysValAsnLeuThrLeuGlyAsp 460
Qy 1439 GACAAAGCCGGCTCTCACAGACCCCTGATTTCTGTTCTGACAGACCCAGCCTCG 1498
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuileSerValProAspArgAspProAlaSer 480
Qy 1499 CCTTTAAGGATGGCAACAGCGCCTGATCTCCGTTGGCTGCTGGCCATATTTGTCAC 1558
Db 481 ProLeuArgMetAlaAsnSerAlaLeuileSerValGlyCysLeuAlailePheValThr 500
Qy 1559 GTGATCTCCCTCTTGGTGTAACAAAAACACAGGAATACACCCCAATAGAAAATAGTCCT 1618
Db 501 VallieSerLeuLeuValTyrlslyshislyslsGluTyrcysProIleGluAsnSerPro 520
Qy 1619 GCGAATGTGGTCAGAAAGCGCCTGAGTGCTCTTCTCAACCGTGCAAGCCGTGTC 1678
Db 521 GlyAsnValValArgSerlyslsGlyLeuSerValPheLeuAsnArgAlaValPhe 540
Qy 1679 TTCCCGGAACACCAAGGAAGGATCGCTACTCAAAAAACCAAGAAATTTAAAGGAGTTTCT 1738
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

RESULT 3
ID ABB78200
XX ABB78200 standard; protein; 560 AA.
AC ABB78200;
DT 25-NOV-2002 (first entry)
DE Amino acid sequence of human HGFIN.
XX Human; cell differentiation; white blood cell; bone marrow cell;
KW haematopoietic growth factor inducible neurokin-1; HGFIN;
KW progenitor proliferation; acute myeloid leukemia; non-Hodgkin's disease;
KW acute lymphocytic leukemia; chronic myeloid leukemia;
KW chronic lymphocytic leukemia; Hodgkin's disease.
XX Homo sapiens.
OS
XX Key
XX Location/Qualifiers
XX 1. .485
XX Domain /note= "extracellular domain"
XX FT Misc-difference 256 /note= "Leu encoded by CC"
XX FT 485. .508
XX FT Domain /note= "transmembrane domain"
XX FT 509. .560
XX FT Domain /note= "intracellular domain"
XX FT
XX WO200262947-A2.
XX
XX 15-AUG-2002.
XX
XX 20-OCT-2001; 2001WO-US050204.
XX
XX 20-OCT-2000; 2000US-0241881P.
XX
XX (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.
XX
XX Pranela R;
XX
XX WPI; 2002-657531/70.
XX N-PSDB; ABQ78551.
XX
XX Hematopoietic growth factor inducible neurokin-1 type polypeptide and
XX PT polynucleotide for treating a disease associated with abnormal bone
XX PT marrow cell differentiation or proliferation, e.g. leukemia.
XX
XX Claim 12; Page 123-125; 125pp; English.
XX
XX The present sequence represents haematopoietic growth factor inducible

```

```

CC neurokin-1 type (HGFIN) polypeptide. HGFIN induces white blood cell
CC differentiation and inhibits progenitor proliferation. HGFIN polypeptides
CC and polynucleotides are useful for treating a disease associated with
CC abnormal bone marrow cell differentiation or proliferation, especially
CC acute myeloid leukemia, acute lymphocytic leukemia, chronic myeloid
CC leukemia, chronic lymphocytic leukemia, Hodgkin's and non-Hodgkin's
CC disease
XX
SQ Sequence 560 AA;

Alignment Scores:
Pred. No.: 4.82e-295 Length: 560
Score: 2999.00 Matches: 559
Percent Similarity: 99.82% Conservative: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 61.84% Indels: 1
DB: 5 Gaps: 0

US-10-039-272-1 (1-2661) x ABB78200 (1-560)
Qy 60 ATGGAATGCTCTACTATTCTCTGGGATTTCCTGGCTCTGGCTGCAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrcysPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20
Qy 120 GCCGCCAAACGATTTCATGATGTGTGGCAATGAAGACCTTCCTCTTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrcysMetArgGlu 40
Qy 180 CACAATCAATTAATGGCTGTCTCTGATGAAATGATGGAATGAAATCTTACCA 239
Db 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrcysPro 60
Qy 240 GTGTGGAAGCGGGAGACATGAGGTGGAAACTCTCGAAGGGAGCGCTGTGAGGGCG 299
Db 61 VallirPlyAsrGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla 80
Qy 300 GTCCTGACAGTACTACACGCCCTCGTGGCTCAATATTAACATTTGCGGTGACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
Qy 360 ATATTCCCTAGATGCCAAAGGAAGATGCCAATGCCAATAGTATGATGAGAGACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrcysAsnGly 120
Qy 420 AGAAATGAGCTGGTTTATCTCTGATCCATATGTTTACCACTGGACAGCATGGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrcysValTyrcysValAsnTrpThrAlaTrpSerGlu 140
Qy 480 GACAGTACGGGGAAAATGGCACCGCAAGCCATCATACGCTCTTCCCTGATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisHisAsnValPheProAspGlyLys 160
Qy 540 CCTTTTCTCACACCCCGGATGGAGAGATGGAATTCATCTAGCTCTTCCACACACTT 599
Db 161 ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrcysValPheHisThrLeu 180
Qy 600 GGTCAAGTATTTCCAGAAATGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCAAT 659
Db 181 GlyGlnTyrcysPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
Qy 660 GTGACACTTGGCCCTCAACTCATGGAAGTACTGTCTACAGAAAGCATGAGCGGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrcysArgHisGlyArgAlaTyrcys 220
Qy 720 GTTCCCATCGCAAGTGAAGATGTGACGTGGTAAACAGATCAGATTCCTGTGTTTGTG 779
Db 221 ValProIleAlaGlnValLysAspValTyrcysValValThrAspGlnIleProValPheVal 240
Qy 780 ACTATGTTTCCAGAAAGCAATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
Qy 839 ATTATGTTTGATGCTCTGATTCATGATCTAGCCACTTCTCAATATTATTCTACCATTAAC 898

```

```

Db      261  IIEMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY      899  TACAGTGGAGCTTCGGGATATACCTGGCCCTGTTTGTTCACCAATACATCTGTAAT 958
Db      281  TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY      959  CACAGTATGTCTCAATGGAACTTCAGCCTTAACCTCACTGTGAAGCTCAGACCA 1018
Db      301  HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY      1019  GGACCTTGTCCGCCCGCCACACACACAGACCTTCAAAACCCACCCCTTCTTTAGGA 1078
Db      321  GlyProCysProProProProProProProProProProProProProProSerLeuGly 340
QY      1079  CCTGCTGGTGACAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACCTGCCAGATTAC 1138
Db      341  ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
QY      1139  AGATATGCCCACTTTCAAGCCACCATCAAAATGTAGAGGGAATCTTAGAGCTTAACATC 1198
Db      361  ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
QY      1199  ATCCAGATGACAGAGCTCTGATGCCGTGCATGCCCTGAAAGCTCCCTAATAGACTTT 1258
Db      381  IleGlnMetThrAspValLeuMetProValProTyrProGluSerSerLeuIleAspPhe 400
QY      1259  GTCGTGACCTGCCAAGGAGCATCCACGAGGCTGTGTACCATCTTCGACCCACC 1318
Db      401  ValValThrCysGlnGlySerIleProThrGluValCysThrIleSerAspProThr 420
QY      1319  TGCAGATCACCCAGAAACACAGCTGCGAGCCTGTGGATGTGGATGGATGATGCTGTG 1378
Db      421  CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY      1379  ACTGTGACAGAACTTCAATGGTCTGGGACGTACTGTGAACCTCACCTCGGGGAT 1438
Db      441  ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460
QY      1439  GACACAGCTGGCTCTCACAGACACCTGATTTCTGTCTCTGACAGAGCCAGCCTCG 1498
Db      461  AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
QY      1499  CCTTAAAGGATGGCAACAGTCCCTGATCCCTGTGGCTGTGGCTGGCCATATTTGTACT 1558
Db      481  ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY      1559  GTGATCTCCCTCTTGGTGTACAAAACACAAAGAAATACACCCCAATAGAAATAGTCCT 1618
Db      501  ValIleSerLeuLeuValTyrLysLysHisLysLysLysLysLysLysLysLysLys 520
QY      1619  GGGATGTGTGTCAGAACCAAGGCTGAGTGTCTTCTCAACCGTGCAGAAAGCGTGTTC 1678
Db      521  GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY      1679  TTCGCGGAAACACAGGAAAGGATCCGCTACTCAAAACCAAGAAATTTAAGGAGTTCT 1738
Db      541  PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

```

## RESULT 4

ABB74961

ID ABB74961 standard; protein; 560 AA.

XX ABB74961;

XX AC

XX DT

XX 01-MAY-2002 (first entry)

XX DE

Human lung tumour L528S protein sequence SEQ ID NO:225.

XX KW

Human; lung tumour; lung cancer; cytostatic; immunostimulant; vaccine; immune response.

XX KW

XX OS

Homo sapiens.

XX

PN

WO200200174-A2.

XX

PD

03-JAN-2002.

XX

PF

28-JUN-2001; 2001WO-US021065.

XX

PR

28-JUN-2000; 2000US-00606421.

PR

02-AUG-2000; 2000US-00630940.

PR

21-AUG-2000; 2000US-00643597.

PR

15-SEP-2000; 2000US-00662786.

PR

09-OCT-2000; 2000US-00685696.

PR

12-DEC-2000; 2000US-00735705.

PR

07-MAY-2001; 2001US-00850716.

XX

PA

(CORI-) CORIXA CORP.

XX

PI

Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;

PI

McNeill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;

PI

Vedwick TS, Carter D, Watanabe Y, Peckham DW;

XX

DR

WPI; 2002-090513/12.

XX

PT

Polynucleotides encoding lung tumor polypeptides, useful for treating

PT

lung cancer or stimulating an immune response.

XX

PS

Disclosure; Page 281-283; 374pp; English.

XX

CC

The present invention describes human lung tumour proteins. Human lung

CC

tumor proteins and polynucleotides have cytostatic and immunostimulant

CC

activities, and can be used in vaccine production. Compositions

CC

comprising the lung tumour proteins, polynucleotides, antibodies, fusion

CC

proteins, T cell populations, or antigen presenting cells that express

CC

the lung tumour proteins are useful for treating lung cancer or

CC

stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to

CC

ABL75070 represent sequences used in the exemplification of the present

CC

invention

XX

SQ

Sequence 560 AA;

Alignment Scores:

Pred. No.: 4.82e-295

Score: 2999.00

Length: 560

Matches: 559

Percent Similarity: 99.82%

Conservative: 0

Best Local Similarity: 99.82%

Mismatch: 1

Query Match: 61.84%

Indels: 1

DB: 5

Gaps: 0

US-10-039-272-1 (1-2661) x ABB74961 (1-560)

QY

60

ATGGAATGCTCTACTATTTCCTGGGATTTCTGCTCTGCTGCAGATTGCCACTTGAT 119

Db

1

MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuLeuAlaAlaArgLeuProLeuAsp 20

QY

120

GCCGCCAAGCATTCATGATGCTGGCAATGAAAGACCTTCTGCTTACATGAGGAG 179

Db

21

AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40

QY

180

CACAAATCAATTAATGGCTGCTTCTGATGAAATGACTGGAAATGAAATCTTACCCA 239

Db

41

HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60

QY

240

GTGTGGAAGCGGGAGACATGAGTGGAAAACTCCTGGAAGGAGCGCTGTGCAAGCG 299

Db

61

ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla 80

QY

300

GTCCTGACCAAGTACTACACAGCCCTCGTGGGCTCAAATATAACATTTGGGTGAACCTG 359

Db

81

ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100

QY

360

ATATTCCTTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAGAACTGC 419

Db

101

IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120

Qy	420	AGAAATGAGGTGGTGTATCTGCTGATCCATATGTTTACAACTGGACAGATGTCAGAG	479	Qy	1499	CCTTTAAGGATGGCAACAGATGCGCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACT	1558
Db	121	ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu	140	Db	481	ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr	500
Qy	480	GACAGTACGGGGAATGGACCGCCAAAGCCATCATACGCTCTCCCTGATGGGAAA	539	Qy	1559	GTGATCTCCCTCTTGGTGTACAAAAACACAAGGAATACAAACCAATAGAAAATAGTCT	1618
Db	141	AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys	160	Db	501	ValIleSerLeuLeuValTyrLysHisLysGluTyrAsnProIleGluAsnSerPro	520
Qy	540	CCTTTTCTCCACCCCGGATGGAGAAGATGGAATTCATCTACGTCTTCACACACTT	599	Qy	1619	GGGAATGTGTGTCAGAACCAAGGCTGAGTCTCTTCTCAACCGTGCAAAAGCCGTGTC	1678
Db	161	ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu	180	Db	521	GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe	540
Qy	600	GGTCAAGTATTCAGAAATGGACGATGTCAGTACAGATGTCGTGCAACACACCAAT	659	Qy	1679	TTCGCGGAAACCAAGAAAGATCCGCTACTCAAAACCAAGAAATTTAAAGGATCTTCT	1738
Db	181	GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn	200	Db	541	PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer	560
Qy	660	GTGACACTTGGGCTCAACTCATGGAAGTACTGTCTACAGAGACATGGACGGCAVAT	719	RESULT 5			
Db	201	ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr	220	ABP61881			
Qy	720	GTCCCATCGCAACAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCCTGTGTGTG	779	ID	ABP61881	standard; protein; 560 AA.	
Db	221	ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal	240	XX	ABP61881;		
Qy	780	ACTATGTTCCAGAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC	838	XX	07-OCT-2002	(first entry)	
Db	241	ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro	260	DE	Human lung cancer associated protein sequence SEQ ID NO:225.		
Qy	839	ATTATGTTTGTGCTCTGATTCATGATCCTAGCCACTTCTCAATTAATTTACCAATTAAC	898	XX	Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.		
Db	261	IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn	280	XX	Homo sapiens.		
Qy	899	TACAAGTGAGCTTCGGGATATACTGCGCTGTTTGTTCACCAATCATACTGTGAAT	958	OS			
Db	281	TyrLysTyrSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn	300	XX	WO200247534-A2.		
Qy	959	CACAGTATGTGCTCAATGGAACCTTCAGCTTAACCTCAGTGTGAAAGTCGACGACCA	1018	PN	20-JUN-2002.		
Db	301	HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro	320	XX	30-NOV-2001; 2001WO-US047576.		
Qy	1019	GGACCTTGTCCGCGCCAGCCACACACACACACCTTCAAAACCCACCTCTTTAGGA	1078	XX	12-DEC-2000; 2000US-00735705.		
Db	321	GlyProCysProProProProProProProProProProProProProProSerLeuGly	340	PR	07-MAY-2001; 2001US-00850716.		
Qy	1079	CCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACCTCCAGATTAAC	1138	PR	28-JUN-2001; 2001US-00897778.		
Db	341	ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn	360	XX	(CORI-) CORIXA CORP.		
Qy	1139	AGATATGGCCACTTTCAGCCACCATCACAATTTAGAGGGAATCTTAGAGTTTAAATC	1198	XX	Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;		
Db	361	ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle	380	PI	McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;		
Qy	1199	ATCCAGATGACAGACGCTCTGATGCCGCTGCTGAGTGTGGAAGCTCCCTAAATAGACTTT	1258	PI	Carter D, Watanabe Y, Peckham DW, Cai F, Foy TM;		
Db	381	IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe	400	XX	WPI; 2002-583465/62.		
Qy	1259	GTCGTGACCTGCCAAGGAGATTCACCGAGGAGTCTGTACCATCATTTCTGACCCCAAC	1318	DR	Novel lung carcinoma polynucleotide sequences and polypeptides encoded by		
Db	401	ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr	420	XX	the polynucleotides, useful in pharmaceutical compositions such as		
Qy	1319	TGGAGATACCCAGACACAGCTCTGAGCCCTGCTGATGTGGATGTGATGTGCTGCTG	1378	XX	vaccines and as markers to indicate the presence of lung cancer.		
Db	421	CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu	440	XX	Example 1; Page 289-290; 381pp; English.		
Qy	1379	ACTGTGAGACAACTTCAATGGGTCTGGGACGTAATGTGTGAACCTCACCTCGGGAT	1438	XX	The present invention describes isolated human lung carcinoma		
Db	441	ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuLeuGlyAsp	460	CC	polynucleotides (I) and polypeptides (II). (I) and (II) have cytostatic		
Qy	1439	GACACAGCTGGCTCTCACGAGCACCTGATTTCTGCTGACAGACCCAGCTCG	1498	CC	activity, and can be used in gene therapy and in vaccines. Compositions		
Db	461	AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer	480	CC	comprising (I) or (II) can be used for stimulating an immune response in		
				CC	a patient and for treating lung cancer in a patient. Oligonucleotides of		
				CC	(I) can be used for detecting the presence of a cancer in a patient, by		
				CC	obtaining a biological sample from the patient, contacting the biological		
				CC	sample with the oligonucleotides, detecting in the sample, an amount of		
				CC	polynucleotide that hybridises to the oligonucleotide and comparing the		
				CC	amount of polynucleotide that hybridises to the oligonucleotide to a		
				CC	predetermined cut-off value, and determining the presence of a cancer in		
				CC	the patient. (I) and (II) are useful in pharmaceutical compositions, e.g.		
				CC	vaccines. (I) is useful as a marker to indicate the presence or absence		
				CC	of a cancer such as lung cancer. ABQ92145 to ABQ92486 and ABP61866 to		
				CC	ABP61992 represent sequences used in the exemplification of the present		
				XX	invention		
				XX	Sequence 560 AA;		
				XX			



## Alignment Scores:

Pred. No.: 4, 82e-295 Length: 560  
 Score: 2999.00 Matches: 559  
 Percent Similarity: 99.82% Conservativity: 0  
 Best Local Similarity: 99.82% Mismatches: 1  
 Query Match: 61.84% Indels: 1  
 DB: 5 Gaps: 0

US-10-039-272-1 (1-2661) x ABP61881 (1-560)

QY 60 ATGAATGTCTCTACTATTTCTCGGATTTCTGCTCGTGGTGAAGATTGCCACTTGAT 119  
 Db 1 MetGluCysLeuTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20  
 QY 120 GCCGCCAAACGATTTCATGTCTGCTGGCAATGAAGACCTTCTGTTACATGAGGAG 179  
 Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40  
 QY 180 CACAATCAATTAAATGGCTGCTCTCTGATGAAATGACTGGAATGAAAACTCTACCCA 239  
 Db 41 HisAsnGlnLeuAsnGlyTyrSerSerAspGluAsnAspTyrAsnGluLysLeuTyrPro 60  
 QY 240 GTGTGGAAGCGGGAGACATGAGTGGAAAACTCTCGGAAGGAGGCCGTGTGACAGCG 299  
 Db 61 ValTyrLysArgGlyAspMetArgTyrLysAsnSerTyrLysGlyArgValGlnAla 80  
 QY 300 GTCTGTACCAAGTCACTCACAGCCCTCGTGGCTCAAAATATACTATTGGGTGAACCTG 359  
 Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAlaAsnLeu 100  
 QY 360 ATATTCCCTAGATGCCAAAAAGAGATGCCAATGGCAACATAGTCTATGAGAGAACTGC 419  
 Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnLys 120  
 QY 420 AGAAATGAGCGCTGTTATCTGTGATCCATATGTTTACAACTGGACAGCATGTCCAGAG 479  
 Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTyrThrAlaTyrSerGlu 140  
 QY 480 GACAGTACCGGGGAAATGCGACGGCCAAAGCCATCAACGTCCTCCCTGATGGGAAA 539  
 Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160  
 QY 540 CCTTTTCTCCACCAACCCCGATGAGAGATGGAATTTTCATCTACGTCCTCCACACTT 599  
 Db 161 ProPheProHisHisProGlyTyrPargArgTyrPasnPhelleIyrValPheHisThrLeu 180  
 QY 600 GGTCAAGTATTCAGAAATGGACGATGTCAGTGAGATTTCTGTGAACACAGCCAAT 659  
 Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200  
 QY 660 GTGACACTTGGCGCTCAACTCATGGAAGTCACTGTCTACAGAAGACNTGGACGGCATAT 719  
 Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220  
 QY 720 GTTCCCATCGCAACAGTGAAGATGTACGTGTAACACATCAGATTCCTGTGTTGTG 779  
 Db 221 ValProIlealaglnValLysAspValTyrValValThrAspGlnIleProValPheVal 240  
 QY 780 ACTATGTTCCAGAAGAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838  
 Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260  
 QY 839 ATTATGTTTGTATCTGATTCATGATTCCTAGCCACTTCTCAATATTTACATTAAAC 898  
 Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280  
 QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCTCTGTTGTTTCCACCAATCATATCTGAT 958  
 Db 281 TyrLysTyrSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300  
 QY 959 CACACGTATGTCTCAATGAACTTTACGCTTTAAGCTCAGCTGTGAAAGCTGCAGACCA 1018  
 Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaAlaPro 320

QY 1019 GGACCTTGTGCGCACCCGCCACACCCAGACCTTCAAAACCCACCCCTTCTTTAGGA 1078  
 Db 321 GlyProCysProProProProProProProProProProProProProProProProSerLeuGly 340  
 QY 1079 CCTGCTGGTGACAAACCCCTGGAGCTGAGTAGATTCTGATGAAAACTGCCAGATTAAAC 1138  
 Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360  
 QY 1139 AGATATGGCCACTTTCAAGCCACCATCACAAATTCAGAGGAATCTTAGAGGTTAAACATC 1198  
 Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAlaAsnIle 380  
 QY 1199 ATCCAGATGACAGACGTCCTGATGCCGTGCCCTGAAAGCTCCCTAATAAGACTTT 1258  
 Db 381 IleGlnMetThrAspValLeuMetProValProTyrProGluSerSerLeuIleAspPhe 400  
 QY 1259 GTCGTGACCTGCCAAGGGAGATTCACCGAGGTCTGTACCATCATTTTGACCCCAAC 1318  
 Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420  
 QY 1319 TGGCAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCCTG 1378  
 Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440  
 QY 1379 ACTGTGAGACGAACCTTCAATGGTCTGGGACGTACTGTGTGAACCTCACCCCTGGGGAT 1438  
 Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460  
 QY 1439 GACACAGCTGGCTCTCACGACACCCCTGATTTCTGTTCTCTCACAGACCCAGCTCG 1498  
 Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480  
 QY 1499 CCTTTAAGGATGGCAACAGTGCCTGATCCCGTTGGCTGTGGCCATATTGTCACT 1558  
 Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500  
 QY 1559 GTCATCTCCCTCTGTGTGTAACAAAAACAAGGAATACAAACCAATAGAAAAATAGTCCT 1618  
 Db 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520  
 QY 1619 GGAATGTGTCAGAACCAAGCCCTGAGTGTCTTCTCAACCGTGCAGAAAGCGGTTC 1678  
 Db 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540  
 QY 1679 TTCCCGGAAAAACAGAAAAAGGATCCGCTACTCAAAAAACAAGATTTAAAGGATTCT 1738  
 Db 541 PheProGlyAsnGlnLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560  
 RESULT 6  
 ABUS6592  
 ID ABUS6592 standard; protein; 560 AA.  
 XX  
 AC ABUS6592;  
 XX  
 DT 02-APR-2003 (first entry)  
 XX  
 DE Lung cancer-associated polypeptide #185.  
 XX  
 KW Lung cancer-associated polypeptide; cytostatic; emphysema;  
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;  
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200286443-A2.  
 XX  
 PD 31-OCT-2002.  
 XX  
 PF 18-APR-2002; 2002WO-US012476.  
 XX

PR 18-APR-2001; 2001US-0284770P.  
PR 10-MAY-2001; 2001US-0290492P.  
PR 09-NOV-2001; 2001US-0339245P.  
PR 13-NOV-2001; 2001US-0350666P.  
PR 29-NOV-2001; 2001US-0334370P.  
PR 12-APR-2002; 2002US-0372246P.  
XX (EOSB-) EOS BIOTECHNOLOGY INC.  
PA  
XX  
XX  
PI Aziz N, Murray R;  
XX  
XX WPI; 2003-093161/08.  
DR N-PSDB; ABX76321.  
XX  
XX  
PT Detecting a lung cancer-associated transcript in a cell from a patient  
PT for treating lung cancer, by contacting a biological sample from the  
PT patient with a polynucleotide that exhibits increased or decreased  
PT expression in lung cancer.  
XX  
XX  
PS Claim 27; Page 328; 453pp; English.  
XX  
CC The invention relates to a method for detecting a lung cancer-associated  
CC transcript in a cell from a patient, comprising contacting a biological  
CC sample from the patient with a polynucleotide that selectively hybridizes  
CC to a sequence that is at least 80 % identical to a gene that exhibits  
CC increased or decreased expression in lung cancer samples. Lung cancer-  
CC associated polynucleotides and polypeptides are used for identifying a  
CC compound that modulates a lung cancer-associated polypeptide, for  
CC inhibiting proliferation of a lung cancer-associated cell to treat lung  
CC cancer in a patient and for treating a mammal having lung cancer by  
CC administering a modulatory compound identified. The methods are useful  
CC for treating lung cancer, such as small cell lung cancer, non-small cell  
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,  
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,  
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and  
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful  
CC for diagnostic purposes and as targets for screening for therapeutic  
CC compounds that modulate lung cancer, such as antibodies. Sequences  
CC ABU56408-ABU56745 represent lung cancer-associated polypeptides of the  
CC invention  
XX  
SQ Sequence 560 AA;  
  
Alignment Scores:  
Pred. No.: Length: 560  
Score: 2999.00 Matches: 559  
Percent Similarity: 99.82% Conservative: 0  
Best Local Similarity: 99.82% Mismatches: 1  
Query Match: 61.84% Indels: 1  
DB: 6 Gaps: 0  
  
US-10-039-272-1 (1-2661) x ABU56592 (1-560)  
  
QY 60 ATGGAATGCTCTACTATTTCTCTGGGATTTCTGCTCTGCTGCTGCAAGATTGCCACTTGAT 119  
Db 1 MetGluCysLeuTyr-Tyr-PheLeuGlyPheLeuLeuLeuAlaAalaArg-LeuProLeuAsp 20  
QY 120 GCCGCCAACGATTTTCATGATGCTGGGCAATGAAGACCTTCTCTTACATGAGGGAG 179  
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40  
QY 180 CACAATCAATTAATGCTGCTCTCTGATGAATGACTGGAATGAAAACTCTACCCA 239  
Db 41 HisAsnGlnLeuAsnGlyTyrSerSerAspGluAsnAspTyrAsnGluLysLeuTyrPro 60  
QY 240 GTGTGAAGCGGGAGACATGAGGTGGAAAACTCTCTGGAAGGGAGCGCGTGTGAGGGCG 299  
Db 61 ValTyrLysArgGlyAspMetArgTyrLysAsnSerTyrLysGlyGlyArgValGlnAla 80  
QY 300 GTCCTGACACGACTCACCAGCCCTCTGGGCTCAATATACATATACATTTCTGGGTGACCTG 359  
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100

QY 360 ATATTCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGATGC 419  
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120  
QY 420 AGAAATGAGGCTGGTTTATCTGTGATCCATATGTTTACAACTGGACAGCATGCTCAGAG 479  
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTyrPheAlaTyrSerGlu 140  
QY 480 GACAGTCACGGGAAATGGCACCGGCCAAGCCATCATACGCTCTTCCCTGATGGGAAA 539  
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160  
QY 540 CTTTTTCTCACCACCCCGGATGGAGAAGATGGAATTTTCATCTACGTCTTCCACACATT 599  
Db 161 ProPheProHisHisProGlyTyrArgArgTyrAsnPheIleTyrValPheHisThrLeu 180  
QY 600 GGTGAGTATTTCCAGAAATTTGGGACGATGTTTCAGTGGAGATTTCTGTGAACACACGCCAAT 659  
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200  
QY 660 GTGACACTTGGGCTCAACTCATGGAAGTGAAGTCTCTACAGAAGACATGGACGGGCATAT 719  
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220  
QY 720 GTTCCCATCGCACAAAGTGAAGATGTGTACTGTGTAAAGATTCAGATTCCTGTGTGTG 779  
Db 221 ValProIleAlaGlnValLysAspValTyrValThrAspGlnIleProValPheVal 240  
QY 780 ACTATGTTCCAGAGAAGCATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838  
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260  
QY 839 ATTATGTTGTATGCTCTGATTCATGATCTGATCCATGCTCTCTCAATTTTACCATTAAAC 898  
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280  
QY 899 TACAAAGTCGAGCTCGGGGATAAATCTGGCTGTTTGTTCACCAATCATCTACTGTGAT 958  
Db 281 TyrLysTyrSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300  
QY 959 CACAGTATGTGCTCAATGGACCTTCAGCCCTTACCTCAGTGTGAAAGCTGCACACCA 1018  
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaAlaPro 320  
QY 1019 GGACCTTGTGCGCCACCCGCCACACCCACCCAGACCTTCAAAACCCACCCCTTCTTTAGGA 1078  
Db 321 GlyProCysProProProProProProProProProProProProProProProSerLeuGly 340  
QY 1079 CCTGCTGTGACAAACCCCTCGAGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1138  
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360  
QY 1139 AGATATGCCACTTTCAAGCCACCATCAATTTGTAGAGGGAATCTTAGAGTTTAAACATC 1198  
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuValAsnIle 380  
QY 1199 ATCCAGATGACAGACGCTCTGATGCGGTGCGATGCGCTGAAAGCTCCCTTAATAGACTTT 1258  
Db 381 IleGlnMetThrAspValLeuMetProValProTyrProGluSerSerLeuIleAspPhe 400  
QY 1259 GTCGTGACCTGCCAAGGGAGCATTCGCCAGGAGTGTGTACCATCTATTCTGACCCACC 1318  
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420  
QY 1319 TCGGAGATCACCCAGACACAGCTCTGCAGCCCTGTGGATGTGGATGTGGATGTGTCTGTG 1378  
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440  
QY 1379 ACTGTGAGACGAACTTCAATGGGCTGTGGAGCGTACTGTGTGAACCTCACCTCGGGGAT 1438  
Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460  
QY 1439 GACACAAGCCTGGCTCTCAGAGACCCCTGATTTCTGTCTTCGACAGACCCAGCCTCG 1498

Db 461 AspThrSerLeuAlaLeuThrSerThrLeuLeuSerValProAspArgAspProIleSer 480  
QY 1439 CTTTAAAGATGGCAACAGTCCCTGATCCCGTGGCTGCTGGCCATATTGTCACT 1558  
Db 481 ProLeuArgMetAlaAsnSerAlaLeuLeuSerValGlyCysLeuAlaIlePheValThr 500  
QY 1559 GTGATCTCCCTCTGGTGTACAAAAACACACAGGAAATACACCCCAATAGAAATAGTCT 1618  
Db 501 ValIleSerLeuLeuValThrLysLysHisLysGluThrAsnProIleGluAsnSerPro 520  
QY 1619 GGAATGTGTCAGAGCAAAAGGCTGAGTGTCTTTCTCAACCGTGCACAAAGCCGTGTC 1678  
Db 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540  
QY 1679 TTCCCGGGAACACAGGAAAGATCCGCTACTCAAAACCAAGATTTAAAGGATTTCT 1738  
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560  
RESULT 7  
ID ABG72962 standard; protein; 560 AA.  
XX AC ABG72962;  
XX DT 08-APR-2003 (first entry)  
XX DE Human osteoactivin homologue, nmb.  
XX KW Human; osteoactivin; osteopathic; antiinflammatory; vaccine;  
KW bone cell differentiation stimulator; gene therapy; bone formation;  
KW osteoporosis; periodontal disease; ectopic bone formation; osteopetrosis;  
KW bone disorder; Osteogenesis; enzyme.  
XX OS Homo sapiens.  
XX PN US2002151486-A1.  
XX PD 17-OCT-2002.  
XX PF 30-AUG-2001; 2001US-00943075.  
XX PR 30-AUG-2000; 2000US-0229006P.  
XX PA (POPO/) POPOFF S N.  
XX PA (SAFA/) SAFADI F F.  
XX PA (OWEN/) OWEN T A.  
XX PA (SMOC/) SMOCK S L.  
XX PI Popoff SN, Safadi FF, Owen TA, Smock SL;  
XX WPI; 2003-182528/18.  
XX PT Novel isolated osteoactivin protein and gene encoding the protein, useful  
PT for stimulating bone differentiation and for treating bone disorders  
PT including osteoporosis and periodontal disease.  
XX PS Claim 24; Fig 2B; 38pp; English.  
XX CC The invention describes an isolated and substantially pure osteoactivin  
CC protein (I) that stimulates bone cell differentiation. A therapeutic  
CC composition comprising (I), the polynucleotide (I) encoding (I) or a  
CC biologically active fragment of (I) is useful for stimulating bone  
CC formation in a mammal. The composition is also useful for treating a bone  
CC disorder including osteoporosis and periodontal disease. A second  
CC therapeutic composition comprising an anti-(I)-antibody or an agent that  
CC inhibits osteoactivin-mediated bone formation is useful for inhibiting  
CC formation in a mammal. The second composition is also useful for treating  
CC ectopic bone formation and osteopetrosis. (I) and (II) are also useful  
CC for developing novel therapeutic compositions for bone disorders, and for  
CC stimulating osteogenesis. The polynucleotide is also useful in gene  
CC therapy. This is the amino acid sequence of human nmb, a homologue of the  
CC rat osteoactivin of the invention used in the creation of antibodies for

CC detection of nmb and osteoactivin  
XX SQ Sequence 560 AA;  
Alignment Scores:  
Pred. No.: 4.82e-295 Length: 560  
Score: 2999.00 Matches: 559  
Percent Similarity: 99.82% Conservative: 0  
Best Local Similarity: 99.82% Mismatches: 1  
Query Match: 61.84% Indels: 1  
DB: 6 Gaps: 0  
US-10-039-272-1 (1-2661) x ABG72962 (1-560)  
QY 60 ATGGAATGCTCTACTATTCTCTGGGATTTCTCTGGCTGCTGCAAGATTGCCACTTGAT 119  
Db 1 MetGluCysLeuTyTyPheLeuGlyPheLeuLeuLeuAlaAlaArgLeuProLeuAsp 20  
QY 120 GCCGCCAAACGATTTCATGTCTGGGCAATGAAAGACCTTCTGCTTACATGAGGAG 179  
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyMetArgGlu 40  
QY 180 CACAATCAATTAAATGCTGTCTTCTCATGAAATGACTGGAATGAAACTCTACCCA 239  
Db 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTy-Pro 60  
QY 240 GTGTGAAGCGGGGAGACATGAGTGGAAAAAATCTCTGGAAGGAGGCCCTGTGCAGCG 299  
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla 80  
QY 300 GTCTGTACAGTAGTACTCACAGCCCTCGTGGGTCAAAATATAACATTTGCGGTGAACCTG 359  
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100  
QY 360 ATATTCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAGAACTGC 419  
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyGluLysAsnCys 120  
QY 420 AGAAATGAGGCTGTTTATCTGCTGATCCATATGTTTACAACTGGACAGCATGCTCAGAG 479  
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyTyValTyAsnTrpThrAlaTrpSerGlu 140  
QY 480 GACAGTAGCGGGGAAAATGCGCCGCAAGCCATCATACCTCTTCCCTGATGGAAA 539  
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160  
QY 540 CTTTTTCCACACCCCGATGAGAGAGATGAAATTCATCTACGTCCTCCACACTT 599  
Db 161 ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyValPheHisThrLeu 180  
QY 600 GGTCACTATTTCAGAAAATTTGGAGCATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAT 659  
Db 181 GlyGlnTyPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200  
QY 660 GTGACACTGGGCTCAACTCATGGAAGTCACTGTCTACAGAGACATGACGCGGATAT 719  
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyArgArgHisGlyArgAlaTy 220  
QY 720 GTTCCCATCGCAAGTGAAGATGTAGTGGTGAACAGATCAGATTCTGTGTTGTG 779  
Db 221 ValProIleAlaGlnValLysAspValTyValValThrAspGlnIleProValPheVal 240  
QY 780 ACTATGTTCCAGAAAGCAATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838  
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260  
QY 839 ATTATGTTTGATGCTCTGATTCATGATCCTAGCCACTTCTCAATTATTCTACCAATAC 898  
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTySerThrIleAsn 280  
QY 899 TACAGTGGAGCTTCGGGATATATCTGCGCTGTTGTTTCCACCAATCATACTGTGAAT 958  
Db 281 TyLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300

QY 959 CACAGTATGCTCAATGGAACTTCAGCCCTTAACCTCACTGTGAGAGCTGCAGACCA 1018  
Db 301 HisThrTyValLeuAsnGlyThrPheSerLeuAsnLeuThrValValAlaAlaPro 320  
QY 1019 GGACCTTCGCGCCACCCACCCACCCAGACCTTCACAAACCCACCCCTCTTTAGGA 1078  
Db 321 GlyProCysProProProProProProProProProProProProProProPro 340  
QY 1079 CCTGCTGTGACAAACCCCTCGAGCTGAGTAGGATTCCTGATGAAACTGCAGATTAA 1138  
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgileProAspGluAsnCysGlnleAsn 360  
QY 1139 AGATATGCCACTTTCAGCCACCATCACAATTCAGAGGAGTCTAGAGGTTAAACATC 1198  
Db 361 ArgTyGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380  
QY 1199 ATCCAGATGACAGACGCTCCTGATGCGGTGCGATGGCTGAAAGCTCCCTAATAGCTTT 1258  
Db 381 IleGlnMetThrAspValLeuMetProValProValProValProValProValProVal 400  
QY 1259 GTCGTGACTGCCAAGGAGAGATTCACCGAGAGTCTGTACATCATTTCTGACCCACC 1318  
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleSerAspProThr 420  
QY 1319 TGCAGATCACCAGAACACAGTCTGCAGCCCTGCGATGGATGGATGAGATGTCGTCTG 1378  
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440  
QY 1379 ACTGTGAGACGAACCTCAATGGGTCTGGGAGTACTGTGTGAACCTCACCTCGGGGAT 1438  
Db 441 ThrValArgThrPheAsnGlySerGlyThrTyCysValAsnLeuThrLeuGlyAsp 460  
QY 1439 GACAAAGCTGTCTCAGAGACCCCTGATTCCTGCTTCTTCCTCAGAGACCCAGCTCG 1498  
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480  
QY 1499 CTTTAAAGATGCGAAACAGTCCCTGATCTCCGTTGGTGTCTGCGCATATTTGTCAT 1558  
Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500  
QY 1559 GTGATCTCCCTCTTGGTGTAACAAAACACAAAGGAATACACCCCAATAGAAATAGTCT 1618  
Db 501 ValIleSerLeuLeuValTyrlslyshislysglytyrAsnProIleGluAsnSerPro 520  
QY 1619 GGAATGTGTGAGAAAGCAAGCCGTGATGTCTTCTCAACCGTGCAAAACCGGTGTC 1678  
Db 521 GlyAsnValValArgSerlysglyLeuSerValPheLeuAsnArgAlaValPhe 540  
QY 1679 TTCCTGGGAACCCAGGAAGGATCCGCTACTCAAAACCAAGAAATTTAAAGGAGTTCT 1738  
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

RESULT 8  
ID ABU70852  
XX ABU70852 standard; protein; 560 AA.  
AC ABU70852;  
XX 10-JUN-2003 (first entry)  
DT Human adipocyte Selected Interacting domain, SID, #483.  
XX Human; prey; adipocyte; SID; selected interacting domain; anorectic;  
KW antidiabetic; protein-protein interaction; diabetes;  
KW yeast 2-hybrid assay; metabolic disorder; obesity.  
XX Homo sapiens.  
XX WO200286122-A2.  
XX 31-OCT-2002.  
XX

PF 14-MAR-2002; 2002WO-EP003768.  
XX  
PR 14-MAR-2001; 2001US-0275734P.  
XX  
FA (HYBR-) HYBRIGENICS.  
XX  
PI Legrain P, Daviet L;  
XX  
DR WPI; 2003-103412/09.  
DR N-FSDB; ACAS7396.  
XX  
PT New complex between two interacting proteins in adipocyte cells, useful  
PT for identifying selected interacting domains that modulate protein  
PT interactions, or for preventing or treating metabolic disorders such as  
PT obesity or diabetes.  
XX  
PS Claim 6; Page 269-270; 382pp; English.  
XX  
CC The invention relates to a complex between two interacting proteins in  
CC adipocyte cells, given in the specification. The proteins are identified  
CC by selecting a bait protein from a known adipocyte marker and then  
CC performing a yeast 2-hybrid selection to isolate prey proteins encoded by  
CC members of an adipocyte cDNA library. The proteins are designated SID  
CC (RTM) (selected interacting domains) proteins. Also included are a  
CC polynucleotide encoding a polypeptide in the adipocyte cells, a  
CC recombinant host cell expressing at least one of the interacting  
CC polypeptides of the complex, selecting a modulating compound in adipocyte  
CC cells, a SID (RTM) polypeptide comprising any of the 738 amino acid  
CC sequences given in the specification (including its fragment or variant),  
CC a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences  
CC given in the specification (including its fragment or variant), a vector  
CC comprising the SID (RTM) polynucleotide, a recombinant host cell  
CC comprising the vector, a protein chip comprising the polypeptides and a  
CC record comprising all or part of the data, listed in the specification.  
CC The complex, polypeptides, polynucleotides and compounds are useful for  
CC preventing or treating metabolic disorders such as obesity or diabetes.  
CC The polynucleotides are useful as probes or primers. The complex is  
CC particularly useful for identifying selected interacting domains (SID  
CC (RTM)) for screening drugs that modulate the protein interaction, thus  
CC exhibiting the therapeutic effect. The present sequence represents a SID  
CC (prey) protein of the invention  
XX  
SQ Sequence 560 AA;  
Alignment Scores:  
Pred. No.: 4, 82e-295 Length: 560  
Score: 2999.00 Matches: 559  
Percent Similarity: 99.8% Conservative: 0  
Best Local Similarity: 99.8% Mismatches: 1  
Query Match: 61.84% Indels: 1  
DB: 6 Gaps: 0  
US-10-039-272-1 (1-2661) x ABU70852 (1-560)  
QY 60 ATGGAATGTCTTACTATTCTCTGGATTTCTGCTGCTGGTGCAGATGTCACCTGTAT 119  
Db 1 MetGluCysLeuTy-Tyr-PheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20  
QY 120 GCCGCCAAACGATTTTCATGATGCTGGCATGAAGACCTTCTGCTTACATGAGGAG 179  
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaLysMetArgGlu 40  
QY 180 CACAATCAATTAATGCGTGTCTTCTGATGAAATGACTGGAATGAAACCTTACCCA 239  
Db 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyPro 60  
QY 240 GTGTGGAACCGGGGACACATGAGGTGGAAAACTCTCTGGAAGGAGCCGCTGTGAGGCG 299  
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla 80  
QY 300 GTCCCTGACCACTGACTCACCAGCCCTCGTGGGCTCAATATATACATTTTGGGTGAACCTG 359  
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100

QY 360 ATATTTCCTAGATGCCAAAGGAGATGCCCAATGGCAACATAGTCTATGAGAAAGTGC 419  
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120  
QY 420 AGAAATGAGGCTGGTTTATCTGTGATCCATATGTTTACAACTGGACAGCATGTGTGAG 479  
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTyrThrAlaTyrSerGlu 140  
QY 480 GACAGTACGGGGAATATGACCGCGCAAGCAATCATACGTCCTCCCTGTATGGGAAA 539  
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160  
QY 540 CCTTTTCCTCACCACCCCGGATGAGAGATGGAAATTCATCTACGTCCTCCACACACTT 599  
Db 161 ProPheProHisLysProGlyTyrArgGlyTrpArgGlyTrpAsnPheIleTyrValPheHisThrLeu 180  
QY 600 GGTGAGTATTTCCAGAAATTTGGACAGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAAT 659  
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200  
QY 660 GTGACACTTGGGCTCAACTCATGGAAGTGAAGTCTCTACAGAAACATGACGGGCATAT 719  
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220  
QY 720 GTTCCCATCCACAAAGTGAAGATGTGTAGTGTGTACAGATCAGATTCCTGTGTTGTG 779  
Db 221 ValProIleAlaGlnValLysAspValTyrValThrAspGlnIleProValPheVal 240  
QY 780 ACTATGTTCCAGAAAGACGATGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838  
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260  
QY 839 ATTATGTTGATGCTGTGATTCATGATGATGATGATGATGATGATGATGATGATGAT 898  
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280  
QY 899 TACAAGTGGAGCTTCGGGATATACCTGGCTCTGTTGTTTCCACCAATCATCTGTGAAT 958  
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300  
QY 959 CACAGCTATGTGCTCAATGAAACCTTCAGCTTAACTCTGAGAAAGCTGACAGACCA 1018  
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320  
QY 1019 GGAACCTTGGCCGACCGCCACACACCCAGACCTTCAAAACCCACCCCTCTTTAGGA 1078  
Db 321 GlyProCysProProProProProProProProProProProProProProProSerLeuGly 340  
QY 1079 CTGCTGGTGTACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCCAGATTAC 1138  
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360  
QY 1139 AGATATGGCCACTTCAAGCCACCATCACAAATGTAGAGGAATCTTAGAGGTTAATC 1198  
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380  
QY 1199 ATCCAGATGACAGACGCTCTGTATGCGGTGCGCATGCGCTGAAAGCTCCCTTAATAGACTTT 1258  
Db 381 IleGlnMetThrAspValLeuMetProValProValProValProValProValProValPhe 400  
QY 1259 GTGCTGACCTGCCAAGGAGCATTCACCGGAGTCTGTACCATCATATTCTGACCCACC 1318  
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420  
QY 1319 TCGAGATCACCCAGAACACAGTCTGAGCCCTGTGATGTGATGATGATGATGATGATGATG 1378  
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440  
QY 1379 ACTGTGAGACGAACTTCAATGGGTCTGGGACCTACTGTGTGAACCTCACCTGGGGGAT 1438  
Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460

QY 1439 GACACAGACCTGGCTCTCAGAGACCCCTGATTTCTGTTCTGACAGAGACCCAGCCTCG 1498  
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480  
QY 1499 CCTTTAAGGATGCAAAACAGTCCCTGATCTCCGTTGCTGCTTGGCCATATTTGTCAC 1558  
Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500  
QY 1559 GTGATCTCTCTTGGTGTACAAAACACAAAGCAATACAAACCAATAGAAATAGTCCT 1618  
Db 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520  
QY 1619 GCGAATGTGTGAGACCAAGCCTGAGTGTCTTCTCAACCGTGCAAAAGCGTGTTC 1678  
Db 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540  
QY 1679 TTCCTGGGAAACCAAGAAAGGATCCGCTACTCAAAAACCAAGAAATTAAGGAGTTCT 1738  
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560  
RESULT 9  
ADA28315  
ID ADA28315 standard; protein; 560 AA.  
XX  
AC ADA28315;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Human lung tumour L528S protein.  
XX  
KW cancer; lung cancer; gene therapy; vaccine; human;  
KW lung squamous cell carcinoma.  
XX  
OS Homo sapiens.  
XX  
PN US2003064947-A1.  
XX  
PD 03-APR-2003.  
XX  
PF 30-NOV-2001; 2001US-00007700.  
XX  
PR 18-MAR-1998; 98US-00040802.  
PR 27-JUL-1998; 98US-00123912.  
PR 22-DEC-1998; 98US-00221107.  
PR 02-APR-1999; 99US-00285479.  
PR 17-DEC-1999; 99US-00466396.  
PR 30-DEC-1999; 99US-00476496.  
PR 10-JAN-2000; 2000US-00480884.  
PR 22-FEB-2000; 2000US-00510376.  
PR 04-APR-2000; 2000US-00542615.  
PR 28-JUN-2000; 2000US-00606421.  
PR 02-AUG-2000; 2000US-00630940.  
PR 21-AUG-2000; 2000US-00643597.  
PR 15-SEP-2000; 2000US-00662786.  
PR 09-OCT-2000; 2000US-00685696.  
PR 12-DEC-2000; 2000US-00735705.  
PR 07-MAY-2001; 2001US-00850716.  
PR 28-JUN-2001; 2001US-00897778.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Wang T, Wang A, Skeiky YAM, Li SX, Kalos MD, Henderson RA;  
PI McNeill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;  
PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TW;  
XX  
XX WPI; 2003-540798/51.  
XX  
XX New isolated polynucleotides and polypeptides useful for diagnosing,  
PT preventing and/or treating cancer, particularly lung cancer.  
XX  
XX Example 7; Page 194-195; 296pp; English.  
XX  
XX The invention describes isolated polynucleotides and polypeptides useful

for diagnosing, preventing and/or treating cancer, particularly lung cancer. A new isolated polynucleotide comprises: any of the 22 fully defined nucleotide sequences (e.g. 1012, 900 or 2773 bp) given in the specification; complements of the nucleotide sequences cited above; at least 10 contiguous residues of the nucleotide sequences cited above; a sequence that hybridize to any of the nucleotide sequences under highly stringent conditions; a sequence that is at least 75 or 90% identical to the above nucleotide sequences; or degenerate variants of the above nucleotide sequences. The composition and methods are useful in diagnosing, preventing and/or treating cancer, particularly lung cancer, in gene therapy and in vaccines. This is the amino acid sequence encoded by a human lung tumour cDNA isolated from a lung squamous cell carcinoma that may be useful in the diagnosis and treatment of lung cancer and other disorders.

XX Sequence 560 AA;

#### Alignment Scores:

Pred. No.: 4.82e-295 Length: 560  
Score: 2999.00 Matches: 559  
Percent Similarity: 99.82% Conservative: 0  
Best Local Similarity: 99.82% Mismatches: 1  
Query Match: 61.84% Indels: 1  
DB: 7 Gaps: 0

US-10-039-272-1 (1-2661) x ADA28315 (1-560)

QY 60 ATGGAATGCTCTACTATTTCTGGGATTCTGCTCGGCTGCAAGATTGCCACTTGAT 119  
DB 1 MetGluCysLeuTyrPheLeuGlyPheLeuLeuLeuAlaAArgLeuProLeuAsp 20  
QY 120 GCCGCCAAACGATTTCATGATGTCTGGCAATGAAGACCTCTGCTTACATGAGGAG 179  
DB 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40  
QY 180 CACAAATCAATTAATGCTGCTCTGATGAAATGACTGGAATGAAATGCTTACCCA 239  
DB 41 HisAsnGlnLeuAsnGlyTyrSerSerAspGluAsnAspTyrAsnGluLysLeuTyrPro 60  
QY 240 GTGTGGAACGGGAGACATGAGTGGGAAACTCTCTGGAGGAGGCGGTGTCAGCGG 299  
DB 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla 80  
QY 300 GTCCTGACAGTACTCACAGCCCTCTGGCTCAATATAACATTTGCGGTGAACCTG 359  
DB 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100  
QY 360 ATATTCCCTAGATGCCAAAGAGATGCCAATGGCAACATAGTCTATCAGAGAAGCTGC 419  
DB 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120  
QY 420 AGAATGAGGCTGTTTATCTGCTGATCCATATGTTTACAACGTGACAGCATGTCAGAG 479  
DB 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140  
QY 480 GACAGTACGGGGAATGGACCGCCCAAGCCCATATAAGCTCTCCCTGATCGGAAA 539  
DB 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160  
QY 540 CCTTTTCTCACCCCGGATGGAGAAGATGGAATTTTCATCTACGCTCTCCACACACTT 599  
DB 161 ProPheProHisHisProGlyTyrArgArgTrpAsnPheIleTyrValPheHisThrLeu 180  
QY 600 GGTGAGTATTTCCAGAAATGGACGATGTTTCAGTGAGAGTTTCTGTGACACACCAAT 659  
DB 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200  
QY 660 GTGACACTGGGCTCAACTCATGGAAGTACTGTCTACAGAGAAGCATGGACGGCATAT 719  
DB 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220  
QY 720 GTTCCCATCGCAACAAAGTGAAGATGTGTACGTGGTGAACAGATCAGATTCTGTGTTGG 779

DB 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240  
QY 780 ACTATGTTCCAGAAAGACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838  
DB 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGlnThrPheLeuLysAspLeuPro 260  
QY 839 ATTATGTTTGTATGTCCTGATTGATTCATGATCCCTAGCCACTTCTCTCAATTATTCTACCATTAA 898  
DB 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280  
QY 899 TACAAGTGGAGCTTCGGGGATATATCTAGCTGCTGTTGTTTCCCAATCATCTGTTGAT 958  
DB 281 TyrIysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300  
QY 959 CACAGCTATGTCTCAATGGAACCTTCAGCTTAACTCAGCTTAACTCAGCTGGAAGCTGCAGACCA 1018  
DB 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320  
QY 1019 GGACCTTGTCCGCCACCGCCACACCCAGACCTTCAAAACCCACCCCTTCTTTAGGA 1078  
DB 321 GlyProCysProProProProProProProProProProProProProProSerLeuGly 340  
QY 1079 CCTGCTGTTGACAAACCCCTGGAGCTGAGTAGGATTCTGATGAAACTGCCAGATTAA 1138  
DB 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360  
QY 1139 AGATATGGCCACTTTCAGGCCACCATCACAAATCTGAGAGGAAATCTTAGAGTTAAACATC 1198  
DB 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380  
QY 1199 ATCCAGATGACAGACGTCCTGATCCGCTGCGCTGCGCTGAAAGCTCCCTTAATAGCTTT 1258  
DB 381 IleGlnMetThrAspValLeuMetProValProValProValProValProValProVal 400  
QY 1259 GTCGTGACCTCCCAAGGAGGACATTCACGAGGAGTCTGTACATCATCTTCTGACCCACC 1318  
DB 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420  
QY 1319 TGCCAGATACCCAGAACACACAGTCTGACGCCCTGTGGATGTGGATGAGATGTCTGCTGTG 1378  
DB 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440  
QY 1379 ACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCCCTGGGGAT 1438  
DB 441 ThrValArgAspThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460  
QY 1439 GACACAAGCTGGCTCTCACGAGCACCTGATTTCTGTTCTGACAGAGACCCAGCCCTCG 1498  
DB 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480  
QY 1499 CCTTTAAGGATGGCAACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCAC 1558  
DB 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500  
QY 1559 GTGATCTCCCTCTTGGTGTACAAAACACAAAGGAATACACCCCAATGAAAATAGTCT 1618  
DB 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520  
QY 1619 GGGATGTGTGACAGCAAGGCTGATGCTTCTTCAACCGTGAAGACCGCTGTC 1678  
DB 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540  
QY 1679 TTCCCGGAAACCAAGAAAGATCCGCTACTCAAAAACCAAGAAATTTAAAGGAGTCTCT 1738  
DB 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560  
RESULT 10  
ADD78235  
ID ADD78235 standard; protein; 572 AA.  
XX  
AC ADD78235;  
XX  
DT 29-JAN-2004 (first entry)

XX Human CGDD-16.  
 DE Anabolic; Hypertensive; Respiratory; Anti-HIV; Antiallergic;  
 KW Neuroprotective; Nootropic; Antianemic; Antiarteriosclerotic;  
 KW Antiinflammatory; Ophthalmologic; Muscular; Hepatotropic;  
 KW Neuroprotective; Antiasthmatic; Anticonvulsant; Virucide; Antibacterial;  
 KW Fungicide; Antiparasitic; Protozoacide; Antihelminthic; Cytostatic;  
 KW Cerebroprotective; Antiparkinsonian; Antiporiatic; Antigout;  
 KW Antidiabetic; Antiarthritic; Antirheumatic; Osteopathic; Gene therapy;  
 KW human; cell growth; cell differentiation; cell death; CGDD;  
 KW cell proliferative disorder; cancer; developmental disorder;  
 KW neurological disorder; autoimmune disorder; inflammatory disorder;  
 KW infection; reproductive disorder.  
 XX Homo sapiens.  
 OS  
 XX WO2003077875-A2.  
 PN  
 XX 25-SEP-2003.  
 XX  
 XX 14-MAR-2003; 2003WO-US008310.  
 XX  
 XX 15-MAR-2002; 2002US-0364494P.  
 PR 29-MAR-2002; 2002US-0369129P.  
 PR 12-APR-2002; 2002US-0372511P.  
 XX  
 XX (INCY-) INCYTE GENOMICS INC.  
 PA  
 XX Kable AE, Tran UK, Hafalia AJA, Burford N, Honchell CD;  
 PI Lehr-Mason PM, Duggan BM, Ramkumar J, Griffin JA, Richardson TW;  
 PI Elliott VS, Jiang X, Jackson AA, Marquis JP, Chawla NK, Khare R;  
 PI Becha SD, Lee SY, Swarnakar A, Yue H, Warren BA, Baughn MR, Lal PG;  
 PI Lee S, Ho A, Gandhi AR, Yao MG;  
 XX WPI; 2003-779081/73.  
 DR N-PSDB; ADD78274.  
 DR  
 XX  
 PT New polypeptides and polynucleotides associated with cell growth,  
 PT differentiation and death, useful for diagnosing, treating or preventing  
 PT e.g. developmental, neurological, autoimmune, inflammatory or  
 PT reproductive disorders.  
 PT  
 XX Claim 1; SEQ ID NO 16; 320pp; English.  
 XX  
 CC The present invention relates to novel human proteins (I; ADD78220-  
 CC ADD78258) and their coding sequences (II; ADD78259-ADD78297), which are  
 CC associated with cell growth, differentiation and death, referred to as  
 CC CGDD-n proteins, where n is a number from 1 to 39. The CGDD proteins and  
 CC their coding sequences are useful for diagnosing, treating or preventing  
 CC cell proliferative disorders (e.g. cirrhosis, hepatitis,  
 CC arteriosclerosis, psoriasis, primary thrombocytopenia) or cancers (e.g.  
 CC adenocarcinoma, sarcoma or cancers of the bone, bone marrow, brain,  
 CC breast, colon, kidney, liver, lung or uterus), developmental disorders  
 CC (e.g. renal tubular acidosis, Becker muscular dystrophy, gonadal  
 CC dysgenesis, hypothyroidism or seizures), neurological disorders (e.g.  
 CC Pick's disease, cataract, epilepsy, ischemic cerebrovascular disease,  
 CC stroke, Alzheimer's disease, Parkinson's disease or dementia),  
 CC autoimmune/inflammatory disorders (e.g. AIDS, allergies, anemia, asthma,  
 CC diabetes mellitus, bronchitis, osteoporosis, osteoarthritis, rheumatoid  
 CC arthritis, contact dermatitis or gout), viral, bacterial, fungal,  
 CC parasitic, protozoan or helminthic infections, reproductive disorders  
 CC (e.g. infertility, ecopic pregnancy, premature ovarian failure, delayed  
 CC puberty or prolatitis) or disorders of the placenta (e.g. preclampsia,  
 CC chorocarcinoma, placenta previa, placental or maternal floor infarction  
 CC or chronic villitis).  
 XX  
 SQ Sequence 572 AA;

## Alignment Scores:

Pred. No.: 4.16e-293 Length: 572  
 Score: 2980.00 Matches: 558  
 Percent Similarity: 97.73% Conservative: 1

Best Local Similarity: 97.55% Mismatches: 1  
 Query Match: 61.44% Indels: 13  
 DB: 7 Gaps: 1

US-10-039-272-1 (1-2661) x ADD78235 (1-572)

QY 60 ATGGAATGCTCTACTATTTCCTGGATTCTGCTGGCTGCAAGATTCCCACTTGAT 119  
 Db |||||||  
 QY 1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaAsgLeuProLeuAsp 20  
 Db |||||||  
 QY 120 GCCGCCAAGCATTTCTATGATGCTCGGCAATCAAGACCTTCTGCTTACATGAGGAG 179  
 Db |||||||  
 QY 21 AlaAlaYsArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40  
 Db |||||||  
 QY 180 CACAATCAATTAAATGGCTGCTCTTCTGATGAAATGACTGGAATGAAAACTCTACCCA 239  
 Db |||||||  
 QY 41 HisAsnGlnLeuAsnGlyTyrSerSerAspGluAsnAspTrpAsnGluLeuLeuTyrPro 60  
 Db |||||||  
 QY 240 GTGTGGAAGCGGGAGACATGAGGTGGAATAAATCTCTGGAGGAGCGCGTGCAGCG 299  
 Db |||||||  
 QY 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla 80  
 Db |||||||  
 QY 300 GTCCTGACCATGACTCACCGACCTCGTGGCTCAATATATAATTAACATTTCGGTGAACCTG 359  
 Db |||||||  
 QY 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnLeuThrPheAlaValAsnLeu 100  
 Db |||||||  
 QY 360 ATATTCCTAGATGCCAAAAAGAGATGCCAATGGCAACATAGTCTATGAGAACTGC 419  
 Db |||||||  
 QY 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnLeuValTyrGluLysAsnCys 120  
 Db |||||||  
 QY 420 AGAAATGAGCTGTTTATCTGCTGATCCATATGTTTACAACTGGACAGCATGCTCAGAG 479  
 Db |||||||  
 QY 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140  
 Db |||||||  
 QY 480 GACAGTACCGGGAANAATGGCAGCGCCCAAGCAATCAACGCTTCCCTGATGGGAAA 539  
 Db |||||||  
 QY 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160  
 Db |||||||  
 QY 540 CTTTTCCTCACCACCCCGGATGGAGAAGATGGAATTCATCTACGCTTCCACACACTT 599  
 Db |||||||  
 QY 161 ProPheProHisHisProGlyTyrArgArgTrpAsnPhelleTyrValPheHisThrLeu 180  
 Db |||||||  
 QY 600 GGTCAGTATTTCCAGAAATGGGACGATGTCAGTGAGAGATTTCTGTGAACACACGCCAAT 659  
 Db |||||||  
 QY 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200  
 Db |||||||  
 QY 660 GTCACTTTGGGCTCAACTCATCGGAAGTCACTGTCTACAGAAGACATGACGGGCATAT 719  
 Db |||||||  
 QY 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220  
 Db |||||||  
 QY 720 GTTCCCATCGCACAAAGTGAAGATGTGCTAGTGAACAGATCAGATTCTCTGTGTTTGTG 779  
 Db |||||||  
 QY 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnLeuProValPheVal 240  
 Db |||||||  
 QY 780 ACTATGTTCCAGAAAGACGATCGAAATTCATCGAGAAACCTTCCC-AAAGATCTCCCC 838  
 Db |||||||  
 QY 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260  
 Db |||||||  
 QY 839 ATTATGTTGATGCTCTGATTCATGATCTAGCCACTTCTCAATTTCTTACCATTAAC 898  
 Db |||||||  
 QY 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280  
 Db |||||||  
 QY 899 TACAAGTGGAGCTTCGGGGATAATCTGGCTGTTTGTGTTTCCACCACATCATCTGTGAAT 958  
 Db |||||||  
 QY 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300  
 Db |||||||  
 QY 959 CACACGTATGTGCTCAATCGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGACCA 1018  
 Db |||||||  
 QY 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320  
 Db |||||||  
 QY 1019 GGACCTTGCCGCCACCGCCACCCACCCAGACCTTCAAAACCCACCCCTCTTTA--- 1075  
 Db |||||||  
 QY 321 GlyProCysProProProProProProProProProProSerLysProThrProSerLeuAla 340  
 Db |||||||



1076 -----GGACCTGCTGGTGAACCCCTGGAG 1102  
341 ThrThrLeuLysSerTyAspSerAsnThrProGlyProAlaGlyAspAsnProLeuGlu 360  
1103 CTGAGTAGATTCTGTAGTAAACATGCGAGATTAAACAGATATGGCATTTCACGACACC 1162  
361 LeuSerArgIleProAspGluAsnGlnIleAsnArgTyrGlyHisPheGlnAlaThr 380  
1163 ATCACAATTGTAGAGGATCTTACAGATTAAACATCATCCAGATGACAGAGCTCCTCATG 1222  
381 IleThrIleValGluGlyIleLeuGluValAsnIleIleGlnMetThrAspValLeuMet 400  
1223 CCGGTGCCATGGCTGAAAGCTCCCTAATAGACTTTGTCTGTCACCTGCCAAGGAGCATT 1282  
401 ProValProTyrProGluSerSerLeuIleAspPheValValThrCysGlnGlySerIle 420  
1283 CCACGAGGTTGTACCATCATTTCTGACCCCATCGGAGATACCCAGAAACACAGTC 1342  
421 ProThrGluValCysThrIleIleSerAspProThrCysGluIleThrGlnAsnThrVal 440  
1343 TGCAGCCCTGTGATGTGATGATGCTGTCTGCTGACTGTGAGACGAACTTCAATGGG 1402  
441 CysSerProValAspValAspGluMetCysLeuLeuThrValArgValThrPheAsnGly 460  
1403 TCTGGGACGCTACTGTGTGAACCTCAACCTGGGGGATGACACAAGCCTGGCTCTCACGAGC 1462  
461 SerGlyThrTyThrCysValAsnLeuThrLeuGlyAspThrSerLeuAlaLeuThrSer 480  
1463 ACCTGATTTCTGCTGACAGAGACCCAGCCTGCCTTTAAGATGGCAACAGTGCC 1522  
481 ThrLeuIleSerValProAspArgAspProAlaSerProLeuArgMetAlaAsnSerAla 500  
1523 CTGATCTCCGTGGCTGGTGGCCATATTGTCACTGTGATCTCCCTCTGTGTACAAA 1582  
501 LeuIleSerValGlyCysLeuAlaIlePheValThrValIleSerLeuValTyrIlys 520  
1583 AAACACAGGAATACAAACCAATAGAAAATAGTCTGGGAATGTGGTCAAGCAAAAGGC 1642  
521 LysHisGlyGluTyrAsnProIleGluAsnSerProGlyAsnValValArgSerIlysgly 540  
1643 CTGAGTGTCTTTCTCAACCGTGCAAAAGCGGTGTTCTCCCGGAAACACAGAAAGGAT 1702  
541 LeuSerValPheLeuAsnArgAlaIlyAlaValPhePheProGlyAsnGlnGlnIlyAsp 560  
1703 CCCTACTCAAAACCAAGATTTAAGGAGTTTCT 1738  
561 ProLeuLeuLysAsnGluGluPheLysGlyValSer 572  
RESULT 11  
AAU83612  
ID AAU83612 standard; protein; 572 AA.  
XX AC AAU83612;  
XX DT 08-MAY-2002 (first entry)  
XX DE Human PRO protein, Seq ID No 42.  
XX KW Human; secreted protein; PRO; tumour; lung cancer; colon cancer;  
KW breast cancer; prostate tumour; rectal tumour; liver tumour;  
KW pericyte cell proliferation; chondrocyte cell proliferation;  
KW tumour necrosis factor-alpha.  
XX OS Homo sapiens.  
XX FN WO200208288-A2.  
XX PD 31-JAN-2002.  
XX PF 29-JUN-2001; 2001WO-US021066.  
XX PR 20-JUL-2000; 2000US-0219556P.



QY 120 GCGCCAAACGATTTCATGTCTGCTGGCAATGAACACCTTCTGCTTACATCAGGGAG 179  
Db 21 AlalaIysArgPheHisAspValLeuGlyAsnGluArgProSerAlaIatyrMetArgGlu 40  
QY 180 CACAATCAATTAATGGCTGCTCTTCATGAAATGACTGGAATGAAACTCTACCCA 239  
Db 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60  
QY 240 GTGTGGAGGGGAGACATGAGTGGGAAACCTCTGGAAAGGAGCGCTGTCCAGCG 299  
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla 80  
QY 300 GTCTGTACCGTACTCACAGCCCTCGTGGGTCMAATATATAACATTTGGGTGAACTG 359  
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100  
QY 360 ATATTCCCTAGATCCCAAGGAAGATGCCAATGGCAACATAGTCTATGAGAACTGC 419  
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120  
QY 420 AGAATGAGCGTGTATTCTGTGATCCATATGTTTACAACTGGAAGATGTCAGAG 479  
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140  
QY 480 GACAGTCACGGGGAAATGGCACCGCCAAAGCCATCATACGCTTCCCTGATGGGAA 539  
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160  
QY 540 CTTTTCCTCACCACCCGGATGGAGAAGATGGAATTTTCATCTACGCTTCCACACACTT 599  
Db 161 PropheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180  
QY 600 GGTCAAGTATTCAGAAATGGAGCAGATGTTCACTGAGAGTTCTGTGAAACACAGCCAAT 659  
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200  
QY 660 GTACACTTCGGCTCAACTCATGGAAGTACTGTCTACAGAGACATGACGGGCATAT 719  
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220  
QY 720 GTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAACAGATCAGATTCTCTGTTTGTG 779  
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240  
QY 780 ACTATGTTCCAGAAAGACGATCGAAATTCATCCGAAACCTTCCC-AAAGATCTCCCC 838  
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260  
QY 839 ATTATGTTTCATGTCCTGATTTCATGATCCTAGCCACTTCCTCAATATTCTACCATTAAC 898  
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280  
QY 899 TACAAGTGGAGCTTCGGGGATAACTCTGGCTGTTTGTTCCTCCAACTCATCTGTGAAT 958  
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300  
QY 959 CACACGATATGTCATCAATGAACCTTCAGCCTTAACCTCATCTGTGAAAGCTGAGACCA 1018  
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320  
QY 1019 GGACCTTGTCGGCCACCGCCACACACAGCCCTTCAAAACCCACCCCTCTTTA--- 1075  
Db 321 GlyProCysProProProProProProProProProProProProProProProProPro 340  
QY 1076 -----GGACCTGCTGTGTGACAAACCCCTCGGAG 1102  
Db 341 ThrThrLeuLysSerTyrAspSerAsnThrProGlyProThrGlyAspAsnProLeuGlu 360  
QY 1103 CTGAGTAGGATTCCTGATGAAACTGCCAGATTAAACAGATATGCGCACTTTTCAAGCCACC 1162  
Db 361 LeuSerArgIleProAspGluAsnCysGlnIleAsnArgTyrGlyHisPheGlnAlaThr 380  
QY 1163 ATCACAATTGTAGGGGAATCTTAGAGTTTAAATCATCATCCAGATGACAGAGCTCTGATG 1222

Db 381 IleThrIleValGluGlyIleLeuGluValAsnIleIleGlnMetThrAspValLeuMet 400  
QY 1223 CCGGTGCGCATGGCTGAAAGCTCCCTAATAGACTTTTGTGTCGACCTGCGCAAGGAGCAT 1282  
Db 401 ProValProTrpProGluSerSerLeuIleAspPheValValThrCysGlnGlySerIle 420  
QY 1283 CCCACGGAGGTCTGTACCATCTTCTGACCCACCTCGCAGATCACCCGAAACACAGTC 1342  
Db 421 ProThrGluValCysThrIleSerAspProThrCysGluIleThrGlnAsnThrVal 440  
QY 1343 TGACAGCCCTGTGATGTGGATGAGATGTCTGCTGACTGTGAGACGAAACCTTCAATGG 1402  
Db 441 CysSerProValAspValAspGluMetCysLeuLeuThrValArgArgThrPheAsnGly 460  
QY 1403 TCTGGGACGTAAGTGTGGAACCTCACCTCGGGGATGACACAGCCCTGGCTCTCACGAG 1462  
Db 461 SerGlyThrTyrCysValAsnLeuThrLeuGlyAspAspThrSerLeuAlaLeuThrSer 480  
QY 1463 ACCCTGATTTCTGTCTCTGACAGACCCAGCCTCGCCTTTAAGGATGGCAAAACAGTGCC 1522  
Db 481 ThrLeuIleSerValProAspArgAspProAlaSerProLeuArgMetAlaAsnSerAla 500  
QY 1523 CTGATCTCCGCTTGGCTTGGCCATATTTGTCACTGTGATCTCCCTTGGTGTCAAA 1582  
Db 501 LeuIleSerValGlyCysLeuAlaIlePheValThrValIleSerLeuLeuValTyrLys 520  
QY 1583 AAACACAAAGGAATACACCCCAATAGAAAATAGTCTGGGAATGTGTCAGAAAGCAAGGC 1642  
Db 521 LysHisLysGluTyrAsnProIleGluAsnSerProGlyAsnValValArgSerLysGly 540  
QY 1643 CTGAGTGTCTTTCTCAACCGTCGAAAAGCCGTGTCTTCCCGGAAACCCAGGAAAGGAT 1702  
Db 541 LeuSerValPheLeuAsnArgAlaLysAlaValPhePheProGlyAsnGlnGluLysAsp 560  
QY 1703 CCCTACTCAAAAACCAAGAAATTAAGAGATTCT 1738  
Db 561 ProLeuLysAsnGlnGluPheLysGlyValSer 572  
RESULT 12  
ABU80759  
ID ABU80759 standard; protein; 572 AA.  
XX  
AC ABU80759;  
XX  
DT 23-JUN-2003 (first entry)  
XX  
DE Human PRO polypeptide #21.  
XX  
KW Human; PRO polypeptide; secreted and transmembrane protein;  
KW anti-PRO antibody; diagnostic assay; gene expression; tumour; cytostatic.  
XX  
OS Homo sapiens.  
XX  
PN US2003036635-A1.  
XX  
PD 20-FEB-2003.  
XX  
PF 28-AUG-2002; 2002US-00230163.  
XX  
PR 25-JUL-2000; 2000US-0220638P.  
PR 01-JUN-2001; 2001WO-US017800.  
PR 29-JUN-2001; 2001WO-US021066.  
PR 09-APR-2002; 2002US-00119480.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;  
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;  
XX  
DR WPI; 2003-342045/32.  
DR N-PSDB; ACA66861.  
XX

One hundred and twenty two nucleic acids encoding PRO polypeptides, useful for the manufacture of a medicament for diagnosing or treating tumor.

Claim 11; Fig 42; 314pp; English.

The present invention relates to the isolation of novel human PRO polypeptides, and the polynucleotide sequences encoding them. The PRO polypeptides are secreted and transmembrane proteins. The PRO polypeptides and polynucleotides are useful for preparing a medicament useful in the diagnosis and treatment of tumours. Anti-PRO antibodies are useful in diagnostic assays for PRO, by detecting its expression in specific cells, tissues or serum, and for affinity purification of PRO from recombinant cell culture or natural sources. ABU0739-ABU0860 represent the human PRO polypeptides of the invention. Note: The sequence data for this patent was obtained in electronic format directly from the USPTO web site at [seqdata.uspto.gov/psipdsIDentry.html](http://seqdata.uspto.gov/psipdsIDentry.html)

Sequence 572 AA;

Alignment Scores:  
Pred. No.: 5,26e-293 Length: 572  
Score: 2979.00 Matches: 558  
Percent Similarity: 97.55% Conservative: 0  
Best Local Similarity: 97.55% Mismatches: 2  
Query Match: 61.42% Indels: 13  
DB: 6 Gaps: 1

US-10-039-272-1 (1-2661) x ABU0759 (1-572)

```
QY 60 ATGGAATGCTCTACTATTCTCTGGGATTTCTGCTCTGGCTGCAAGATTGCCACTTGAT 119
DB 1 MetGluCysLeuTyr-Tyr-PheLeuGlyPheLeuLeuLeuAlaAArgLeuProLeuAsp 20
QY 120 GCCGCCAAACGATTTCATGATGTGCTGGCGCAATGAAGACCTTCCTGCTACATGAGGAG 179
DB 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAAATCAATTAATGGCTGCTCTCTGATGAAATGACTGGAATGAAATCTTACCCA 239
DB 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGlnLysLeuTyrPro 60
QY 240 GTGTGGAAGCGGGAGACATGAGGTGGAAATCTCCTCGAAGGAGCGGTGTGCGAGCG 299
DB 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla 80
QY 300 GTCCTGACAGTGACTACAGCCCTCGTGGCTCAATATATAACATTTGCGGTGAACCTG 359
DB 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
QY 360 ATATTCCCTAGATGCCAAAGGAAGATGCCAATGCCAATAGTCTATGAGAAGAACTGC 419
DB 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGlnLysAsnCys 120
QY 420 AGAAATGAGCGTGTGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAG 479
DB 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACAGTACGGGGGAAAATGGACCGCCAAAGCCATCATACGCTCTCCCTGATGGGAAA 539
DB 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
QY 540 CTTTTTCTCACACCCCGGATGGAGAGATGGAATTTCTATCTACGTCTTCACACACTT 599
DB 161 ProPheProHisPheProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCAAGTATTTCCAGAAATPGGACCATGTTTCAGTCAGAGATTTCTGTGAACACAGCCAAT 659
DB 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGGCTCAACTCATGAGTGAAGTACTGTCTACAGAGACATGGACGGCATAT 719
DB 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
```

```
QY 720 GTTCCCATCGCACAAAGTAAAGATGCTGTAACGTGTAAACAGATCAGATTCCTGTGTTGTG 779
DB 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAAGACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCC 838
DB 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTGATGTCCTGATTATCATGATCCTAGACCATCTCCATTAATTCTACCATTAAC 898
DB 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGGATAAATACTGGCCCTGTTTGTTCACCAATCATACTGTGAAT 958
DB 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTCTCAATGGAACTTCAGCCCTTAACCTCACTGTGAAGAGTCGACGACCA 1018
DB 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCCGCCACCGCCACACACCCAGACCTTCAAAACCCACCCCTTCTTTA--- 1075
DB 321 GlyProCysProProProProProProProProProProProProProProProSerLeuAla 340
QY 1076 -----GGACCTGTGCTGTGACAAACCCCTCGAG 1102
DB 341 ThrThrLeuLysSerTyrAspSerAsnThrProGlyProThrGlyAspAsnProLeuGlu 360
QY 1103 CTGAGTAGGATTCCTGATGAAAACTGCCAGATTAACAGATATGCCACTTTCAAGCCACC 1162
DB 361 LeuSerArgIleProAspGluAsnCysGlnIleAsnArgTyrGlyHisPheGlnAlaThr 380
QY 1163 ATCAAAATTGAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGAGCTCCTGATG 1222
DB 381 IleThrIleValGluGlyIleLeuGluValAsnIleIleGlnMetThrAspValLeuMet 400
QY 1223 CCGGTGCCATGGCTGAAAGCTCCCTAATAGACTTTGTGTGTCGCTGCAAGGGAGCATT 1282
DB 401 ProValProTrpProGluSerSerLeuIleAspPheValValThrCysGlnGlySerIle 420
QY 1283 CCCACGGAGGTCTGTACCATCATTTCTGACCCCTGCGAGATCACCAGAACACAGTC 1342
DB 421 ProThrGluValCysThrIleIleSerAspProThrCysGluIleThrGlnAsnThrVal 440
QY 1343 TGCAGCCCTGTGGATGTGGATGAGTGTCTGCTGACTGTGAGACGAACTTCAATGGG 1402
DB 441 CysSerProValAspValAspGluMetCysLeuLeuThrValArgArgThrPheAsnGly 460
QY 1403 TCTGGGACGTACTGTGTGAACCTCACCCCTGGGGATGACAAAGCTGCTCTCACGAGC 1462
DB 461 SerGlyThrTyrCysValAsnLeuThrLeuGlyAspAspThrSerLeuAlaLeuThrSer 480
QY 1463 ACCCTGATTTCTGTCCTGACAGAGACCCGCTCGCTTAAAGATGCCAACAGTCGCC 1522
DB 481 ThrLeuIleSerValProAspArgAspProAlaSerProLeuArgMetAlaAsnSerAla 500
QY 1523 CTGATCTCCGTTGGCTGTGGCCATATTGTCACTGTGATCTCCCTCTTGGTGACAAA 1582
DB 501 LeuIleSerValGlyCysLeuAlaIlePheValThrValIleSerLeuLeuValTyrLys 520
QY 1583 AAACAACAGGAATACACCCCAATAGAAAATAGTCTCGGAATGTGGTCAAGAGCAAGGC 1642
DB 521 LysHisLysGluTyrAsnProIleGluAsnSerProGlyAsnValValArgSerLysGly 540
QY 1643 CTGAGTGTCTTCTCAACCGTCGCAAAAGCCGTGTTCTTCCCGGGAACACGAAAGGAT 1702
DB 541 LeuSerValPheLeuAsnArgAlaLysAlaValPhePheProGlyAsnGlnGluLysAsp 560
QY 1703 CCGCTACTCAAAAACCAAGAAATTTAAAGAGTTTCT 1738
DB 561 ProLeuLeuLysAsnGlnGluPheLysGlyValSer 572
```

## RESULT 13

ABO33725  
ID ABO33725 standard; protein; 572 AA.

AC ABO33725;

XX 17-SEP-2003 (first entry)

XX Novel human secreted and transmembrane protein PRO9925.

XX Human; secreted and transmembrane protein; PRO; cytostatic;  
XX antiarthritic; osteopathic; gene therapy; TNF-Agonist-Alpha;  
XX chondrocyte stimulator; pericyte stimulator; fibroblast modulator;  
XX pharmaceutical; diagnostic; biosensor; bioindicator; lung tumour;  
XX colon tumour; breast tumour; prostate tumour; rectal tumour;  
XX liver tumour; bone disorder; cartilage disorder; sports injury;  
XX arthritis; wound.

XX Homo sapiens.

XX US2003045687-A1.

XX 06-MAR-2003.

XX 12-AUG-2002; 2002US-00218631.

XX 01-JUN-2001; 2001WO-US017800.

XX 29-JUN-2001; 2001WO-US021066.

XX 09-APR-2002; 2002US-00119480.

XX (GETH ) GENENTECH INC.

XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;

XX Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

XX WI; 2003-512315/48.

XX N-PSDB; ACD68613.

XX New genes, and its encoded secreted and transmembrane polypeptides,  
XX useful for stimulating tumor Necrosis Factor alpha, or chondrocyte or  
XX pericyte proliferation, especially for treating lung tumors, arthritis or  
XX wounds in a mammal.

XX Claim 11; Fig 42; 314pp; English.

XX The invention describes an isolated nucleic acid molecule comprising a  
XX sequence with at least 80% identity to: (a) a nucleotide encoding any of  
XX 122 PRO (secreted and transmembrane) polypeptides whose sequences are  
XX fully defined in the specification; or (b) any of 122 nucleotide  
XX sequences having e.g. 4834, 2504 or 1759 bp fully defined in the  
XX specification; or the full length coding sequence of any these 122  
XX nucleotide sequences. The PRO polypeptides or polynucleotides are useful  
XX as pharmaceuticals, diagnostics, biosensors or bioeffectors. These are  
XX particularly useful for detecting tumors (e.g. lung tumour, colon  
XX tumour, breast tumour, prostate tumour, rectal tumour, or liver tumour)  
XX in a mammal, for stimulating the release of TNF-alpha from human blood,  
XX for stimulating the proliferation or differentiation of chondrocyte  
XX cells, for stimulating proliferation of pericyte cells, or for modulating  
XX normal human dermal fibroblast proliferation. The PRO nucleic acid or  
XX polypeptide is also useful for treating tumors or various bone and/or  
XX cartilage disorders (e.g. sports injuries or arthritis), or wounds. The  
XX PRO polypeptides are useful in drug screening, particularly as targets  
XX for therapeutic intervention in these diseases, and in the diagnostic  
XX determination of the presence of these diseases. The PRO polypeptides are  
XX also useful as molecular weight markers, or for chromosome  
XX identification. The PRO genes are useful as hybridisation probes, or for  
XX screening libraries of human cDNA, genomic DNA or mRNA. The PRO genes may  
XX also be used in gene therapy, particularly for replacing a defective  
XX gene. This is the amino acid sequence of a novel human secreted and  
XX transmembrane PRO polypeptide

XX Sequence 572 AA;

## Alignment Scores:

Pred. No.: 5,26e-293 Length: 572  
Score: 2979.00 Matches: 558  
Percent Similarity: 97.55% Conservative: 0  
Best Local Similarity: 97.55% Mismatches: 2  
Query Match: 61.42% Indels: 13  
DB: 6 Gaps: 1

US-10-039-272-1 (1-2661) x ABO33725 (1-572)

QY	60	ATGGAATGCTCTACTATTTCCTGGGATTTCTGCTGCTGCTGCAAGATTGCCACTTGAT	119
DB	1	MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuLeuAlaAlaArgLeuProLeuAsp	20
QY	120	GCGCCAAACGATTTCATGATGCTGGGCAATCAAGACCTTCTGCTTACATGAGGAG	179
DB	21	AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu	40
QY	180	CACAATCAATTAATGGCTGCTTCTGTGATGAAATGACTGGATGAAAACTCTACCCA	239
DB	41	HisAsnGlnLeuAsnGlyTyrSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro	60
QY	240	GTGTGGAAGCGGGAGACATGAGGTGMAAACTCTCTGGGAAGGAGCGCTGTGCGAGCG	299
DB	61	ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla	80
QY	300	GTCTGACCACTGACTCACCAGCCCTCGTGGGCTCAAAATAACATTTGCGGTGAACCTG	359
DB	81	ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu	100
QY	360	ATATTCCCTAGATGCCAAAGGAGATGCCAATGGCAACATAGTCTATGAGAGAACTGC	419
DB	101	IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys	120
QY	420	AGAAATGAGGCTGTTTATCTGCTGATCCATATGTTTACAACTGGACAGCATGTCAGAG	479
DB	121	ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnIleThrAlaIlePheSerGlu	140
QY	480	GACAGTCACGGGAAAATGGCACCGGCCAAAGCCATCAACCTCTTCCTGTATGGGAAA	539
DB	141	AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys	160
QY	540	CTTTTCTCCACACCGCGGATGGAGAAGATGGAATTCATCTACGCTTCCACACACTT	599
DB	161	ProPheProHisHisProGlyTyrArgArgTrpAsnPheIleTyrValPheHisThrLeu	180
QY	600	GGTCAGTATTTCAGAAAATGGGACCATGTTTCAGTGAGAGTTCCTGTGACACAGCAAT	659
DB	181	GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn	200
QY	660	GTGACACTTGGGCTCAACTCATGGAAGTCACTGTCTACAGAAGACATGGACGGCATAT	719
DB	201	ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr	220
QY	720	GTTCCTCATCGCACACTGGAAGATGTCAGTGGTAACTGCTGAGATTCCTGTGTTTGTG	779
DB	221	ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal	240
QY	780	ACTATGTTCCAGAAGAACCATGAAATTCATCGGAGAAACCTTCCC-AAAGATCTCCCC	838
DB	241	ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro	260
QY	839	ATTATGTTTGTCTCTGATTCATGATCTCTAGCCACTTCTCTCAATATTCTTACCATTAAC	898
DB	261	IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn	280
QY	899	TACAAGTCAGCTTCGGGATATACTGGCTGTTGTTTCCACCAATCATACTGTGAAT	958
DB	281	TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn	300
QY	959	CACACGTATGTCTCAATGGAACCTTCAGCTTAACTCACCTGTGAAAGCTGACAGACCA	1018
DB	301	HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro	320



Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120  
QY 420 AGAAATGAGCTGTTTATCTGCTGATCCATATGTTTACAACTGCGACAGCATGCTCAGAG 479  
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140  
QY 480 GACAGTCACCGGGGAAATGSCACCGCGCAAGCCATCATACGTCCTCCCTGATGGGAAA 539  
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160  
QY 540 CCTTTTCTCCACACCCCGGATGGAGAAGATGGAATTTTCATCTACGCTTCCACACACTT 599  
Db 161 ProPheProHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180  
QY 600 GGTCAAGTATTCAGAAATTCGAGACGATGTTCTAGTGAGAGTTCTTGTAACACAGCCAAAT 659  
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200  
QY 660 GTGACACTTGGGCTCAACTCATGGAAGTGAAGTCTCTACAGAGACATGACGGGCATAT 719  
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220  
QY 720 GTTCCCATCCACAAAGTGAAGATGTGTACGTGGTAACAGATCAGATTCCTGTTGTTGTG 779  
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240  
QY 780 ACTATGTTCCAGAAAGACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838  
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260  
QY 839 ATTATGTTTCATGCTCCTGATTGATGCTAGCCAGTCCAGTCCATCTCAATATTACCATTAAC 898  
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280  
QY 899 TACAAGTGGAGCTCGGGGATAATACTGGCGCTGTTGTTTCCCAATCATACTGTGAAT 958  
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300  
QY 959 CACAGTATGCTCAATGAACTTCAGCTTAACCTCACTGAAAGCTGACGACCA 1018  
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320  
QY 1019 GGACCTTGTCGGCCACCGCACACACCCAGACCTTCAAAACCCACCCCTCTTTA--- 1075  
Db 321 GlyProCysProProProProProProProProProProProProProProSerLeuAla 340  
QY 1076 -----GGACCTGCTGTGTGACAAACCCCTGGAG 1102  
Db 341 ThrThrLeuLysSerTyrAspSerAsnThrProGlyProThrGlyAspAsnProLeuGlu 360  
QY 1103 CTGAGTAGGATTCCTGATGAAACTGCGCAGATTACAGATATGCCCATCTTCAAGCCACC 1162  
Db 361 LeuSerArgIleProAspGluAsnCysGlnIleAsnArgTyrGlyHisPheGlnAlaThr 380  
QY 1163 ATCACAATTCAGAGGAACTTAGAGGTTAAATCATCATCCAGATGACAGACGTCCTGATG 1222  
Db 381 IleThrIleValGluGlyIleLeuGluValAsnIleIleGlnMetThrAspValLeuMet 400  
QY 1223 CCGGTGCGATGGGCTGAAAGCTCCCTAATAGACTTTGCTGTGACCTGCCAAGGAGGACTT 1282  
Db 401 ProValProTrpProGluSerSerLeuIleAspPheValValThrCysGlnGlySerIle 420  
QY 1283 CCACGAGGTCGTGATACCATCATTTCTGACCCACCTCGAGATCAACCCAGAACACAGTC 1342  
Db 421 ProThrGluValCysThrIleIleSerAspProThrCysGluIleThrGlnAsnThrVal 440  
QY 1343 TGCAGCCCTGATGATGGATGATGTCGTGATGTCGTGATGTCGACGAACTTCAATGGG 1402  
Db 441 CysSerProValAspValAspGluMetCysLeuLeuThrValArgArgThrPheAsnGly 460  
QY 1403 TCTGGGACGTACTGTGTGAACCTCACCTCGGGGATGACACAAAGCCGTGCTCACGAGC 1462  
Db 461 SerGlyThrTyrCysValAsnLeuThrLeuGlyAspAspThrSerLeuAlaLeuThrSer 480

QY 1463 ACCTGATTTCTGTTCTCTGACAGACCCAGCCTCGCCTTTAAGGATGGCAACAGTGCC 1522  
Db 481 ThrLeuIleSerValProAspArgAspProAlaSerProLeuArgMetAlaAsnSerAla 500  
QY 1523 CTGATCTCCGCTTGGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAA 1582  
Db 501 LeuIleSerValGlyCysLeuAlaIlePheValThrValIleSerLeuLeuValTyrLys 520  
QY 1583 AAACACAAAGGAATACAAACCAACCAATAGTCTCTGGGAATGTGGTCAGAAGCAAGGC 1642  
Db 521 LysHisLysGluTyrAsnProIleGluAsnSerProGlyAsnValValArgSerLysGly 540  
QY 1643 CTGAGTGTCTTTCTCAACCGTCGAAAGCCGTGTTCTCCCGGAAACACGAGAAAGCAT 1702  
Db 541 LeuSerValPheLeuAsnArgAlaLysAlaValPhePheProGlyAsnGlnGluLysAsp 560  
QY 1703 CCCTACTCAAAACCAAGAAATTTAAGAGGATTTCT 1738  
Db 561 ProLeuLysAsnGlnGluPheLysGlyValSer 572  
RESULT 15  
ABJ72248  
ID ABJ72248 standard; protein; 572 AA.  
XX AC ABJ72248;  
XX DT 06-NOV-2003 (first entry)  
XX DE Human PRO9925 protein.  
XX KW PRO; proliferation; pericyte cell; TNF-alpha; blood; chondrocyte;  
XX KW differentiation; dermal fibroblast; tumour; gene therapy; cytostatic.  
XX OS Homo sapiens.  
XX PN US2003050448-A1.  
XX PD 13-MAR-2003.  
XX PF 28-AUG-2002; 2002US-00230414.  
XX PR 01-JUN-2001; 2001WO-US017800.  
XX PR 29-JUN-2001; 2001WO-US021066.  
XX PR 09-APR-2002; 2002US-00119480.  
XX PA (GETH ) GENENTECH INC.  
XX PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;  
XX PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;  
XX DR WPI; 2003-521818/49.  
XX DR N-PSDB; ABT44246.  
XX PT New nucleic acid encoding for a PRO protein, useful for the manufacture  
XX PT of a medicament for diagnosing or treating tumors or for measuring or  
XX PT detecting expression of an associated gene.  
XX PS Claim 11; Fig 42; 315pp; English.  
XX CC The invention relates to a novel isolated nucleic acid encoding a fully  
XX CC defined PRO polypeptide. The molecules of the invention may be useful for  
XX CC stimulating proliferation or gene expression in pericyte cells or the  
XX CC release of TNF-alpha from human blood. Other possible uses include the  
XX CC stimulation or inhibition of chondrocyte proliferation or  
XX CC differentiation, the stimulation of human dermal fibroblast cell  
XX CC proliferation and the detection of the presence of a tumour within a  
XX CC mammal. Furthermore, the nucleic acid may be useful for the manufacture  
XX CC of a medicament for diagnosing or treating a tumour within a mammal or  
XX CC for measuring or detecting the expression of an associated gene, as well  
XX CC as during gene therapy. The current sequence is that of the human PRO  
XX CC protein of the invention

SQ Sequence 572 AA;

## Alignment Scores:

Pred. No.:	5,26e-293	Length:	572
Score:	2979.00	Matches:	558
Percent Similarity:	97.55%	Conservative:	0
Best Local Similarity:	97.55%	Mismatches:	2
Query Match:	61.42%	Indels:	13
DB:	6	Gaps:	1

US-10-039-272-1 (1-2661) x ABU72248 (1-572)

QY	60	ATGGAATGCTCTACTATTTCTGGGATTTCTGCTCTGGCTGCAAGATTCCTCACTTGAT	119
DB	1	MetGluCysLeuTyrPheLeuGlyPheLeuLeuLeuAlaAlaArgLeuProLeuAsp	20
QY	120	GCCGCCAAACGATTTCATGATGTGCTGGCAATGAAAGACCTTCTGCTTACATAGGGAG	179
DB	21	AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu	40
QY	180	CACAATCAATTAATGGCTGCTTCTGATGAAATGACTGGAATGAAAACTCTACCCA	239
DB	41	HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro	60
QY	240	GTGTGGAAGCGGGAGACATGAGGTGGAATACTCTCGAAGGGAGCGCGTGTGAGCGC	299
DB	61	ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla	80
QY	300	GTCCTGACAGTGACTACAGCCCTCTGGGCTCAAAATATAACATTTGGGGTGAACCTG	359
DB	81	ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu	100
QY	360	ATATTCCTAGATGCCAAAGAGAGATGCCAATGCCAATAGTCTATGAGAAGAACATGC	419
DB	101	IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys	120
QY	420	AGAAATGAGGTGGTTTATCTGCTGATCCATATGTTTACAACTGGACAGCATGTGTCAG	479
DB	121	ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu	140
QY	480	GACAGTGACGGGAAATGGACCCGCCAAGCCATCATACGTCTTCCCTGATGGGAAA	539
DB	141	AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys	160
QY	540	CCTTTTCTCACCACCCGGATGGAGATGGATTTCACTAGCTCTCCACACACTT	599
DB	161	ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu	180
QY	600	GCTCAGTATTTCCAGAAATGGACGATGTTTCAGTGAGAGTTTCTGTGAACACACCCAAT	659
DB	181	GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn	200
QY	660	GTGACACTTGGCCCTCACTCATGGAAGTGACTGTCTACAGAAGACATGGAGCGCATAT	719
DB	201	ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr	220
QY	720	GTTCCTCATCGCAAGTGAAGATGTGACGGTGAACAGATTCAGATTCCGTGTGTTCTG	779
DB	221	ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal	240
QY	780	ACTATGTTCCAGAAGACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC	838
DB	241	ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro	260
QY	839	ATTATGTTTGTGCTGCTGATTCATGATCCTACCCACTTCTCCTCAATTATTCTACCATTAAC	898
DB	261	IleMetPheAspValIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn	280
QY	899	TACAAGTCGAGCTTCGGGATAATCTGGCCCTGTTTGTTCACCACAATCATACTCTGAAT	958
DB	281	TyrLysTrpSerPheLysAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn	300
QY	959	CACACGTATGTCTCAATGGAACCTTACGCTTAACTCCTGCTGAGAAAGCTGCAGACCA	1018

DB	301	HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro	320
QY	1019	GGACCTTGTCCGCCACCGCCACACACAGCTTCAAAACCCACCCCTTCTTTA---	1075
DB	321	GlyProCysProProProProProProProProProProProProProProProPro	340
QY	1076	-----GGACCTGCTGGTGACAAACCCCTGTGAG	1102
DB	341	ThrThrLeuLysSerTyrAspSerAsnThrProGlyProGlyAspAsnProLeuGlu	360
QY	1103	CTGAGTAGGATTCCTGATGAAAACCTGCAGATTAAACAGATATGCCCCTTTCAAGCCACC	1162
DB	361	LeuSerArgIleProAspGluAsnCysGlnIleAsnArgTyrGlyHisPheGlnAlaThr	380
QY	1163	ATCACAATTTAGAGGGAATCTTAGAGGTAAACATCATCCAGATGACAGAGCTCTCTGATG	1222
DB	381	IleThrIleValGluGlyIleLeuGluValAsnIleIleGlnMetThrAspValLeuMet	400
QY	1223	CCGCTGCCATGGCCTGAAAGCTCCCTAATAGACTTTGTCTGACCTGCCAAGGGAGCATT	1282
DB	401	ProValProTrpProGluSerSerLeuIleAspPheValValThrCysGlnGlySerIle	420
QY	1283	CCGACGAGGTCTGTACATCTTCTGACCCCACTGCGAGATCACCCAGAACACAGTC	1342
DB	421	ProThrGluValCysThrIleIleSerAspProThrCysGluIleThrGlnAsnThrVal	440
QY	1343	TGCAGCCCTGTGGATGTGGATGATGTCTGCTGCTGCTGAGACGAACTTCAATCGG	1402
DB	441	CysSerProValAspValAspGluMetCysLeuLeuThrValArgArgThrPheAsnGly	460
QY	1403	TCTGGGACGTACTGTGTGAACCTCACCTGGGGGATGACAAAGCCTGGCTCTCACGAGC	1462
DB	461	SerGlyThrTyrCysValAsnLeuThrLeuGlyAspAspThrSerLeuAlaLeuThrSer	480
QY	1463	ACCTGATTTCTGTTCTCTGACAGACCCAGCCTCGCTTAAAGATGGCAACAGTGCC	1522
DB	481	ThrLeuIleSerValProAspArgAspProAlaSerProLeuArgMetAlaAsnSerAla	500
QY	1523	CTGATCTCCGTGGCTGCTGGCCATATTCTCACTGTGTGATCTCCCTCTGTTGTACAAA	1582
DB	501	LeuIleSerValGlyCysLeuAlaIlePheValThrValIleSerLeuLeuValTyrLys	520
QY	1583	AAACACAGGAATACAAACCCCAATAGAAATAGTCTCTGGGAATGTGTCAGAGCAAGGC	1642
DB	521	LysHisLysGluTyrAsnProIleGluAsnSerProGlyAsnValValArgSerLysGly	540
QY	1643	CTGAGTGTCTTCTCAACCGTGCAGAAAGCCGTGTTCTTCCCGGGAAACCCAGAAAAGAT	1702
DB	541	LeuSerValPheLeuAsnArgAlaValPhePheProGlyAsnGlnGluLysAsp	560
QY	1703	CCGCTACTCAAAACCAAGAAATTTAAAGGAGTTTCT	1738
DB	561	ProLeuLeuLysAsnGlnGluPheLysGlyValSer	572

Search completed: September 10, 2004, 15:39:40  
Job time : 236 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.  
OM nucleic - nucleic search, using sw model  
Run on: September 9, 2004, 20:21:02 ; Search time 1483 Seconds  
(without alignments)  
7622.694 Million cell updates/sec  
Title: US-10-039-272-1  
Perfect score: 2661  
Sequence: 1 cggcagaggccagagga.....aaaaaaaaaaaaaaaaaaaaa 2661  
Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 3373863 seqs, 2124099041 residues  
Total number of hits satisfying chosen parameters: 6747726  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_29Jan04.\*  
1: Geneseqn1980s.\*  
2: Geneseqn1990s.\*  
3: Geneseqn2000s.\*  
4: Geneseqn2001as.\*  
5: Geneseqn2001bs.\*  
6: Geneseqn2002s.\*  
7: Geneseqn2003as.\*  
8: Geneseqn2003bs.\*  
9: Geneseqn2003cs.\*  
10: Geneseqn2004s.\*  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				% Query Match		Length DB ID		Description	
Result No.	Score	Query Match	Length DB ID						
1	2661	100.0	2661 6	ABQ78551	Nucleotid				
2	2547.8	95.7	2669 2	AAT69328	Murine me				
3	2547.8	95.7	2669 2	AAT69318	Murine me				
4	2547.8	95.7	2669 3	AAC55715	Human NMB				
5	2547.8	95.7	2669 6	ABQ88185	Human ost				
6	2547.8	95.7	2669 7	ABX76321	Lung can				
7	2505.5	94.2	2728 6	ADD78274	Human CGD				
8	2485	93.4	2683 6	ABK33556	CDNA enco				
9	2485	93.4	2683 7	ACA66861	CDNA enco				
10	2485	93.4	2683 7	ACD68613	Novel hum				
11	2485	93.4	2683 7	ACA68517	Novel hum				
12	2485	93.4	2683 8	ABT44245	Novel hum				
13	2485	93.4	2683 8	ABT44529	Human PRO				
14	2485	93.4	2683 8	ACD82196	Human PRO				
15	2485	93.4	2683 8	ABT43902	Human mem				
16	2485	93.4	2683 8	ADB83531	Novel hum				
17	2485	93.4	2683 8	ADB80637	Novel hum				
18	2485	93.4	2683 8	ADB73178	Novel hum				
19	2485	93.4	2683 8	ADB78260	Novel hum				
20	2485	93.4	2683 9	ADB84908	Human PRO				
21	2485	93.4	2683 9	ADB78014	Novel hum				
22	2485	93.4	2683 9	ADB87080	Human PRO				
23	2485	93.4	2683 9	ADB84662	Human PRO				

24	2485	93.4	2683	9	ADB83777	Novel hum
25	2485	93.4	2683	9	ADB72932	Novel hum
26	2485	93.4	2683	9	ADC36770	Human PRO
27	2485	93.4	2683	9	ADC21760	Human PRO
28	2485	93.4	2683	9	ADC49791	Novel hum
29	2485	93.4	2683	9	ADC48990	Novel hum
30	2485	93.4	2683	9	ADC49507	Novel hum
31	2485	93.4	2683	9	ADC47368	Novel hum
32	2485	93.4	2683	9	ADC47113	Novel hum
33	2485	93.4	2683	9	ADC77988	Novel hum
34	2485	93.4	2683	9	ADD06223	Novel hum
35	2485	93.4	2683	9	ADC77742	Novel hum
36	2485	93.4	2683	9	ADD50705	Novel hum
37	2485	93.4	2683	9	ADD50951	Novel hum
38	2485	93.4	2683	9	ADD50432	Human PRO
39	2485	93.4	2683	9	ADD50186	Human PRO
40	2485	93.4	2683	9	ADD51197	Novel hum
41	2485	93.4	2683	10	ADC48744	Novel hum
42	2485	93.4	2683	10	ADE20915	Novel hum
43	2485	93.4	2683	10	ADE05759	Human PRO
44	2485	93.4	2683	10	ADD74988	Human PRO
45	2485	93.4	2683	10	ADD75734	Novel hum

ALIGNMENTS

RESULT 1		ABQ78551		standard; DNA; 2661 BP.	
ID	ABQ78551	standard; DNA; 2661 BP.			
XX	ABQ78551				
AC	ABQ78551				
XX	25-NOV-2002	(first entry)			
DT	DT				
XX	Nucleotide sequence of human HGFIN.				
DE	Human; cell differentiation; white blood cell; bone marrow cell;				
XX	haematopoietic growth factor inducible neurokin-1; HGFIN;				
KW	progenitor proliferation; acute myeloid leukemia; non-Hodgkin's disease;				
KW	acute lymphocytic leukemia; chronic myeloid leukemia;				
KW	chronic lymphocytic leukemia; Hodgkin's disease; gene; ss.				
XX	Homo sapiens.				
OS	Key	Location/Qualifiers			
XX	CDS	60..1741			
FT		/*tag= a			
FT		/product= "HGFIN"			
FT		/transl_except= (825..826,aa:Leu)			
XX	WO200262947-A2.				
PN	15-AUG-2002.				
XX	20-OCT-2001; 2001WO-US050204.				
PF	20-OCT-2000; 2000US-0241881P.				
XX	(UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.				
PA	Pranela R;				
XX	WPI; 2002-657531/70.				
XX	P-PSDB; ABB78200.				
DR	Hematopoietic growth factor inducible neurokin-1 type polypeptide and				
XX	polynucleotide for treating a disease associated with abnormal bone				
XX	marrow cell differentiation or proliferation, e.g. leukemia.				
PS	Claim 2; Page 121-123; 125pp; English.				
XX	The present sequence encodes human haematopoietic growth factor inducible				
CC					







QY	628	GTTCAGTGAAGATTTCTGTGAACACAGCCAAATGTGACACTTGGGCGCTCAACTCATGGAAG	687
Db	660	GTTCAGTGAAGATTTCTGTGAACACAGCCAAATGTGACACTTGGGCGCTCAACTCATGGAAG	719
QY	688	TGACTCTGTACAGAACAATGACCGGGCATATGTTCCCATCGCAACAAGTGAAGATGTGT	747
Db	720	TGACTCTGTACAGAACAATGACCGGGCATATGTTCCCATCGCAACAAGTGAAGATGTGT	779
QY	748	ACGTGGTAAACAGATCAGATTCCTGTGTGTTGTGACTATGTTCAGAGAACAACCATCGAAATT	807
Db	780	ACGTGGTAAACAGATCAGATTCCTGTGTGTTGTGACTATGTTCAGAGAACAACCATCGAAATT	839
QY	808	CATCCGACGAAACCTTCCC-CAAAAGATCTCCCATTAATGTTTGATGTCCTGATTATCATGATC	866
Db	840	CATCCGACGAAACCTTCCCCTCAAGATCTCCCATTAATGTTTGATGTCCTGATTATCATGATC	899
QY	867	CTAGCCACTTCCTCAATATTCTTACCATTAATCAAGTTGGAGCTTCGGGGATATACTG	926
Db	900	CTAGCCACTTCCTCAATATTCTTACCATTAACTTCAAGTTGGAGCTTCGGGGATATACTG	959
QY	927	GCCTGTTGTTTCCACCAATCATCTGTGAATCACACGTATCTGTCTCAATGGAACCTTCA	986
Db	960	GCCTGTTGTTTCCACCAATCATCTGTGAATCACACGTATGTGTCTCAATGGAACCTTCA	1019
QY	987	GCCTTAACTCTACTGTGAAAGCTGCAGCACAGGACCTTGTCCGCCACGCCACCAACCAC	1046
Db	1020	GCCTTAACTCTACTGTGAAAGCTGCAGCACAGGACCTTGTCCGCCACGCCACCAACCAC	1079
QY	1047	CCAGACCTTCAAAAACCCACCTTCTTTAGGACCTGCTGGTGACAAACCCCTGGAGCTGA	1106
Db	1080	CCAGACCTTCAAAAACCCACCTTCTTTAGGACCTGCTGGTGACAAACCCCTGGAGCTGA	1139
QY	1107	GTAGGATTCCTGATGAAACCTGCCAGATTAAACAGATATGSCCACTTTCAAGCCACAATCA	1166
Db	1140	GTAGGATTCCTGATGAAACCTGCCAGATTAAACAGATATGSCCACTTTCAAGCCACAATCA	1199
QY	1167	CAATTGTAGAGGGAACTTTAGAGGTTAACTATCATCCAGATGACAGACGTCTGTATGCCGG	1226
Db	1200	CAATTGTAGAGGGAACTTTAGAGGTTAACTATCATCCAGATGACAGACGTCTGTATGCCGG	1259
QY	1227	TGCCATGGCCTGAAAGTCCCTTAATAGACTTTGTGTGACCTGCCAAGGGAGGANTCCCA	1286
Db	1260	TGCCATGGCCTGAAAGTCCCTTAATAGACTTTGTGTGACCTGCCAAGGGAGGANTCCCA	1319
QY	1287	CGGAGGTCTGTACCAATCATTTCTGACCCCACTTCGCGAGATCACCCAGAACACAGTCTGCA	1346
Db	1320	CGGAGGTCTGTACCAATCATTTCTGACCCCACTTCGCGAGATCACCCAGAACACAGTCTGCA	1379
QY	1347	GCCCTGTGGATGTGATGAGATGTGTCTGCTGACTGTGAGACGAACTTCAATGGGTCTG	1406
Db	1380	GCCCTGTGGATGTGATGAGATGTGTCTGCTGACTGTGTGAGACGAACTTCAATGGGTCTG	1439
QY	1407	GGACGTACTGTGTGAACTCAACCTTGGGGGATGACACAAGCCTGGCTCTCACAGCACCC	1466
Db	1440	GGACGTACTGTGTGAACTCAACCTTGGGGGATGACACAAGCCTGGCTCTCACAGCACCC	1499
QY	1467	TGATTTCTGTTCTTCACAGAGACCCAGCTCCGCTCTTTAAGATGGCAAAACAGTGCCTGTA	1526
Db	1500	TGATTTCTGTTCTTCACAGAGACCCAGCTCCGCTCTTTAAGATGGCAAAACAGTGCCTGTA	1559
QY	1527	TCTCCGTTGGCTGCTTGGCCATATTGTGCTGATCTCCCTCTTGGTGTACAAAAC	1586
Db	1560	TCTCCGTTGGCTGCTTGGCCATATTGTGCTGATCTCCCTCTTGGTGTACAAAAC	1619
QY	1587	ACAAGGAATACAAACCAATAGAAAATAGTCTCTGGGAATGTGGTCAGAGCAAAAGGCTGA	1646
Db	1620	ACAAGGAATACAAACCAATAGAAAATAGTCTCTGGGAATGTGGTCAGAGCAAAAGGCTGA	1679
QY	1647	GTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCTTCCCGGGAAACACAGGAAAAGGATCCGC	1706
Db	1680	GTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCTTCCCGGGAAACACAGGAAAAGGATCCGC	1739
QY	1707	TACTCAAAAACCAAGAATTTAAAGGAGTTCTTTAAATTTGACACTTGTTCCTGAGCTCA	1766

1740	Db	TACTCAAAAACCAAGAAATTTAAAGGAGTTTCTTAAATTTGCACCTTGTTCCTGAAGCTCA	1799
1767	QY	CTTTTCAGTGCCTATTCATGTGAGATGCTGGAGTGGCTATTAAACCTTTTTTTTCCCTAAAG	1826
1800	Db	CTTTTCAGTGCCTATTCATGTGAGATGCTGGAGTGGCTATTAAACCTTTTTTTTCCCTAAAG	1859
1827	QY	ATTATTGTTTAAATAGATATTGTGGTTTGGGGAAGTTGAAATTTTTTATAGCTTAAATGTC	1886
1860	Db	ATTATTGTTTAAATAGATATTGTGGTTTGGGGAAGTTGAAATTTTTTATAGCTTAAATGTC	1919
1887	QY	TTTTTAGAGATGGGAGAGGGAATTATACCTGCAGGCAGCTTCAGCCATCTGTTGTGAAACCTGAT	1946
1920	Db	TTTTTAGAGATGGGAGAGGGAATTATACCTGCAGGCAGCTTCAGCCATCTGTTGTGAAACCTGAT	1979
1947	QY	AAAAGCAACTTAGCAAGGCTCTCTTTCAATATTTTTTTATGTTTCACTTATAAAGTCTTAG	2006
1980	Db	AAAAGCAACTTAGCAAGGCTCTCTTTCAATATTTTTTTATGTTTCACTTATAAAGTCTTAG	2039
2007	QY	GTAACCTAGTAGGATAGAAAACTGTGTCGCCGAGATGAAGAGAGAACTACTATTGATTA	2066
2040	Db	GTAACCTAGTAGGATAGAAAACTGTGTCGCCGAGATGAAGAGAGAACTACTATTGATTA	2099
2067	QY	GAGCTTAACCCAGGCTTAACCTCAAGAAGAGGCGGATACCTTCAGCTTTCCATGTAACTG	2126
2100	Db	GAGCTTAACCCAGGCTTAACCTCAAGAAGAGGCGGATACCTTCAGCTTTCCATGTAACTG	2159
2127	QY	TATGCATAAAGCCAAATGTAGTCCAGTTTCTAAGATCATGTTTCCAAAGCTAACTGAATCCCA	2186
2160	Db	TATGCATAAAGCCAAATGTAGTCCAGTTTCTAAGATCATGTTTCCAAAGCTAACTGAATCCCA	2219
2187	QY	CTTCAATACACACTCATGAACCTCTGTATGGAACCAATAACAGGCCCAAGCCCTGTGGTATGA	2246
2220	Db	CTTCAATACACACTCATGAACCTCTGTATGGAACCAATAACAGGCCCAAGCCCTGTGGTATGA	2279
2247	QY	TGTGCACACTTGTCTAGACTCAGAAAAAATACTACTCTCATAAATGGGTGGGAGTATTTTG	2306
2280	Db	TGTGCACACTTGTCTAGACTCAGAAAAAATACTACTCTCATAAATGGGTGGGAGTATTTTG	2339
2307	QY	GTGACAACTTACTTTGCTGTGGCTGAGTGAAGGAATGATTTTCATATATTTCATTTATTTCCA	2366
2340	Db	GTGACAACTTACTTTGCTGTGGCTGAGTGAAGGAATGATTTTCATATATTTCATTTATTTCCA	2399
2367	QY	TGGACATTTAGTTAGTGCCTTTTTATATACAGGCAATGATGCTGAGTGACACTCTTTGTGTA	2426
2400	Db	TGGACATTTAGTTAGTGCCTTTTTATATACAGGCAATGATGCTGAGTGACACTCTTTGTGTA	2455
2427	QY	TATTTTCCAAATTTTTGTATAGTGCCTGCACATATTTGAAATCAAAATATTAAAGACTTTCC	2486
2460	Db	TATTTTCCAAATTTTTGTATAGTGCCTGCACATATTTGAAATC-ATATATTAAAGACTTTCC	2511
2487	QY	AAAAATTTGGTCCCTGGTTTTTCATGGCAACTTGATCAGTAGGATTTTCCCTCTGTTTG	2544
2519	Db	AAAGATGAGGTCCTCGGTTTTTCATGGCAACTTGATCAGTAGGATTTTCCCTCTGTTTG	2577
2547	QY	GAACTAAAACCACTTACTATATGTTTAGACAAGACATTTTTTTTTTTTCTCTCTCTGAAAA	2600
2579	Db	TTAAGTAAAACCACTTACTATATGTTTAGACAAGACATTTTTTTTCTCTCTCTCTGAAAA	2636
2607	QY	-AAAAATGGGGAGAGACAAAAAATAAAAAA 2636	
2639	Db	TAAAGTGTGGGAAGAGACAAAAAATAAAAAA 2669	

RESULT 3	
AAT69318	
ID	AAT69318 standard; cDNA; 2669 BP.
XX	
XX	
XX	AAT69318;
XX	
XX	26-FEB-1998 (first entry)
DT	
XX	
XX	
DE	Murine metastatic nucleic acid sequence

XX Mouse; murine; tumour; cancer; metastatic sequence; detection; diagnosis;  
 KW treatment; metastasis; hyperplasia; dysplasia, hypertrophy; screening;  
 KW ss.  
 XX Mus musculus.  
 XX MO9718454-A2.  
 XX 22-MAY-1997.  
 XX 15-NOV-1996; 96WO-US018567.  
 XX 16-NOV-1995; 95US-0006838P.  
 XX 30-JAN-1996; 96US-00594031.  
 XX (THOM/) THOMPSON T.  
 XX Thompson T;  
 XX WPI; 1997-289397/26.  
 XX Identifying tumour metastatic sequences - by introducing transfected  
 PT cells into host mammal and analysing primary and metastatic sequences by  
 PT differential display PCR.  
 XX Disclosure; Fig 12BT; 102pp; English.  
 XX Mouse Urogenital Sinus (UGS) tissue was isolated from 17 day old mouse  
 CC embryos. The UGS cells were infected with retroviruses, cultured and  
 CC implanted under the renal capsule of mice. Reconstructions were harvested  
 CC 5 weeks later, when they showed signs of distress from the tumour burden.  
 CC Metastatised tumours were isolated from a site outside the tumour capsule.  
 CC RNA was isolated from primary tumours and metastases, reverse transcribed  
 CC and subjected to differential display PCR. The sequences were analysed to  
 CC obtain metastatic sequences, e.g. the present sequence. The method can be  
 CC used to detect, diagnose and treat disorders related to metastasis, or  
 CC treat malignant or non-malignant disorders, e.g. hyperplasia, dysplasia  
 CC and hypertrophy. The metastatic sequence can be used to screen a  
 CC biological sample for metastasis, and it or its expression product may  
 CC also be used to treat a metastatic disorder  
 XX Sequence 2669 BP; 752 A; 595 C; 597 G; 725 T; 0 U; 0 Other;  
 SQ  
 Query Match 95.7%; Score 2547.8; DB 2; Length 2669;  
 Best Local Similarity 99.2%; Pred. No. 0;  
 Matches 2591; Conservative 0; Mismatches 17; Indels 3; Gaps 3;  
 28 AACCTTGGTGGCTCGCTCCGTGAGAAATTCAGCATGGAAATGCTCTACTATTTCCTGGGAT 87  
 60 ACCTTGAGTGCTCGCTCCGTGAGAAATTCAGCATGGAAATGCTCTACTATTTCCTGGGAT 119  
 88 TTCTGCTCCGTGGCTGCGAGATTGCCACTTGATGCGGCGAAACGATTTTCATGATGCTGG 147  
 120 TTCTGCTCCGTGGCTGCGAGATTGCCACTTGATGCGGCGAAACGATTTTCATGATGCTGG 179  
 148 GCAATGAAGAACCTTCTGCTTACATGAGGAGACAAATCAATTAATGGCTGCTCTCTG 207  
 180 GCATGAAGAACCTTCTGCTTACATGAGGAGACAAATCAATTAATGGCTGCTCTCTG 239  
 208 ATGAAATGACTGGAATGAAATGAAATGCTTACACAGTGTGGAGCGGGAGACATGAGNGGA 267  
 240 ATGAAATGACTGGAATGAAATGAAATGCTTACACAGTGTGGAGCGGGAGACATGAGNGGA 299  
 268 AAAATCTCTGGAAGGAGCGGTGTCAGCGGCTCTGACACGAGTACTCCACGCGCTCG 327  
 300 AAAATCTCTGGAAGGAGCGGTGTCAGCGGCTCTGACACGAGTACTCCACGCGCTCG 359  
 328 TGGGCTCAAAATATAAATTTGCGGTGAACCTGATATTCCTAGATGTCGCAAAAGAGATG 387  
 360 TGGGCTCAAAATATAAATTTGCGGTGAACCTGATATTCCTAGATGTCGCAAAAGAGATG 419  
 388 CCAATGGCAACATAGTCTATGAGAGAACTGCAGAAATGAGGCTGGTTATCTGCTGATC 447

Db 420 CCAATGGCAACATAGTCTATGAGAGAACTGCAGAAATGAGGCTGGTTATCTGCTGATC 479  
 QY 448 CATATGTTTACAACTGAGCAGCATGTCAGAGACAGTGAACGGGAAATGGCACCGGCC 507  
 Db 480 CATATGTTTACAACTGAGCAGCATGTCAGAGACAGTGAACGGGAAATGGCACCGGCC 539  
 QY 508 AAAGCCCATCAATACGCTCTTCCCTGATGGGAAACCTTTTCTCTACACCCCGGATGGAGAA 567  
 Db 540 AAAGCCCATCAATACGCTCTTCCCTGATGGGAAACCTTTTCTCTACACCCCGGATGGAGAA 599  
 QY 568 GATGGAATTTCACTACGCTCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGAT 627  
 Db 600 GATGGAATTTCACTACGCTCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGAT 659  
 QY 628 GTTCAGTGAGAGTTTCTGTGACACAGCAATGTGACACTTGGGCTCAACTCATGGAAG 687  
 Db 660 GTTCAGTGAGAGTTTCTGTGAACACAGCAATGTGACACTTGGGCTCAACTCATGGAAG 719  
 QY 688 TGACTGTCTACAGAAACATGACGCGGCATATGTTCCCATCGCACAAAGTGAAGATGCTG 747  
 Db 720 TGACTGTCTACAGAAACATGACGCGGCATATGTTCCCATCGCACAAAGTGAAGATGCTG 779  
 QY 748 ACGTGGTAAACAGATCAGATTCCTGTTGTGACTATGTTCCAGAAAGACGATCGAAAT 807  
 Db 780 ACGTGGTAAACAGATCAGATTCCTGTTGTGACTATGTTCCAGAAAGACGATCGAAAT 839  
 QY 808 CATCCGAGCAACCTTCC-CAAGATCTCCCATATGTTTGTATGTCCTGATTCATGATC 866  
 Db 840 CATCCGAGCAACCTTCTCTCAAGATCTCCCATATGTTTGTATGTCCTGATTCATGATC 899  
 QY 867 CTAGCCACTTCTCAATTAATTTCTACCAATTAACAGTGGAGCTTCGGGATAATATCTG 926  
 Db 900 CTAGCCACTTCTCAATTAATTTCTACCAATTAACAGTGGAGCTTCGGGATAATATCTG 959  
 QY 927 GCCTGTTGTTTCCCAATCATATGCTGAATCACACGATGCTGCTCAATGGAACCTTCA 986  
 Db 960 GCCTGTTGTTTCCCAATCATATGCTGAATCACACGATGCTGCTCAATGGAACCTTCA 1019  
 QY 987 GCCTTAACTCACTGTGAAGCTGCAGCACAGGACCTTGTTCGCGCACCGCCACACAC 1046  
 Db 1020 GCCTTAACTCACTGTGAAGCTGCAGCACAGGACCTTGTTCGCGCACCGCCACACAC 1079  
 QY 1047 CCAGACCTTCAAAACCCCACTTCTTAGACCTGCTGTGACAAACCCCTGGAGCTGA 1106  
 Db 1080 CCAGACCTTCAAAACCCCACTTCTTAGACCTGCTGTGACAAACCCCTGGAGCTGA 1139  
 QY 1107 GTAGATTCCTGATGAAACCTGCCAGATTAACAGATATGGCCACTTTCAAGCCACCATCA 1166  
 Db 1140 GTAGATTCCTGATGAAACCTGCCAGATTAACAGATATGGCCACTTTCAAGCCACCATCA 1199  
 QY 1167 CAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGAGTCTGATGCGGG 1226  
 Db 1200 CAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGAGTCTGATGCGGG 1259  
 QY 1227 TGCCATGGCCCTGAAAGCTCCCTCAATAGACTTGTGCTGACCTGCGAAGGGAGGATTCCTCA 1286  
 Db 1260 TGCCATGGCCCTGAAAGCTCCCTCAATAGACTTGTGCTGACCTGCGAAGGGAGGATTCCTCA 1319  
 QY 1287 CGAGAGTCTGTACATCATTTCTGACCCCACTTCGAGATCAACCCAGAAACACAGTCTGCA 1346  
 Db 1320 CGAGAGTCTGTACATCATTTCTGACCCCACTTCGAGATCAACCCAGAAACACAGTCTGCA 1379  
 QY 1347 GCCTGTGAGATGAGATGATGCTGCTGATGCTGAGACGAAACCTTCAATGGGCTG 1406  
 Db 1380 GCCTGTGAGATGAGATGATGCTGCTGATGCTGAGACGAAACCTTCAATGGGCTG 1439  
 QY 1407 GGAAGTACTGTGTGAACCTCACCTCGGGGATGACACAGCCCTGCTTCACAGACACC 1466  
 Db 1440 GGAAGTACTGTGTGAACCTCACCTCGGGGATGACACAGCCCTGCTTCACAGACACC 1499  
 QY 1467 TGAATTTCTGTCTGTACAGAGACCCAGCTTCGCTTTAAGGATGGCAACAGTGCCTGA 1526

Mon Sep 13 10:28:12 2004

us-10-039-272-1.rng

Db 1500 TGATTTCTGTTCTCGACAGACCCAGCGCTCGCTTTTAAGGATGGCAACAGTGGCCCTGA 1559  
Qy 1527 TCTCGTTGGCTGCTTGGCCATATTTGTCATGTGATCTCCCTCTTGTGTGTAACAAAAAC 1586  
Db 1560 TCTCGTTGGCTGCTTGGCCATATTTGTCATGTGATCTCCCTCTTGTGTGTAACAAAAAC 1619  
Qy 1587 ACAGGAATACAAACCAATAGAAATAGTCTGGGAATGTGTCAGAAAGCAAGGSCCTGA 1646  
Db 1620 ACAGGAATACAAACCAATAGAAATAGTCTGGGAATGTGTCAGAAAGCAAGGSCCTGA 1679  
Qy 1647 GTGCTCTTCTAACCGTGCMAAGCGGTGTTCTTCCCGGAAACAGGAAAGGATCCGC 1706  
Db 1680 GTGCTCTTCTAACCGTGCMAAGCGGTGTTCTTCCCGGAAACAGGAAAGGATCCGC 1739  
Qy 1707 TACTCAAAACCAAGAAATTAAGAGATTTCTTAATTTGACCTTGTTCGAAGCTCA 1766  
Db 1740 TACTCAAAACCAAGAAATTAAGAGATTTCTTAATTTGACCTTGTTCGAAGCTCA 1799  
Qy 1767 CTTTTCAGTGCATTCATGTGAGATGCTGCTGGAGTGGCTATTAACTTTTTCCTAAAG 1826  
Db 1800 CTTTTCAGTGCATTCATGTGAGATGCTGCTGGAGTGGCTATTAACTTTTTCCTAAAG 1859  
Qy 1827 ATTATTGTTAAATAGATATGTGTTTGGGGAAGTTGAATTTTATAGGTTAAATGTCA 1886  
Db 1860 ATTATTGTTAAATAGATATGTGTTTGGGGAAGTTGAATTTTATAGGTTAAATGTCA 1919  
Qy 1887 TTTTATAGATGGGAGAGGATATATCTGAGGAGCTTCAGCCATGTTGTGAAGTGTAT 1946  
Db 1920 TTTTATAGATGGGAGAGGATATATCTGAGGAGCTTCAGCCATGTTGTGAAGTGTAT 1979  
Qy 1947 AAAGCAACTAGCAGGCTCTTTTCAATATTTTATGTTTCACTTATAAGTCTTAG 2006  
Db 1980 AAAGCAACTAGCAGGCTCTTTTCAATATTTTATGTTTCACTTATAAGTCTTAG 2039  
Qy 2007 GTAACTAGTAGGATAGAAACACTGTGTCGAGTAAGGAGAGAGCTACTATTGATTA 2066  
Db 2040 GTAACTAGTAGGATAGAAACACTGTGTCGAGTAAGGAGAGAGCTACTATTGATTA 2099  
Qy 2067 GAGCCTAACCCAGGTTAACTGCAAGAGAGCGGGATACCTTCAGCTTCCATGTACTG 2126  
Db 2100 GAGCCTAACCCAGGTTAACTGCAAGAGAGCGGGATACCTTCAGCTTCCATGTACTG 2159  
Qy 2127 TATGCATAAGCCAACTAGTCTAGCTTCTTAAGATCATGTTCCAAAGTAACTAACTCCCA 2186  
Db 2160 TATGCATAAGCCAACTAGTCTAGCTTCTTAAGATCATGTTCCAAAGTAACTAACTCCCA 2219  
Qy 2187 CTTCAATACACACTCATGAACCTCTCTGATGGAACAATAACAGGCCCAAGCCTGTGATGA 2246  
Db 2220 CTTCAATACACACTCATGAACCTCTCTGATGGAACAATAACAGGCCCAAGCCTGTGATGA 2279  
Qy 2247 TGTGCACACTTGTGATGCTCAGAAAAAATCTACTCTCATAAATGGGTGGGAGTATTTTG 2306  
Db 2280 TGTGCACACTTGTGATGCTCAGAAAAAATCTACTCTCATAAATGGGTGGGAGTATTTTG 2339  
Qy 2307 GTGCAACCTACTTGTGCTGCTCAGTGAAGGATGATATTCATATATTCATTTATTTCCA 2366  
Db 2340 GTGCAACCTACTTGTGCTGCTCAGTGAAGGATGATATTCATATATTCATTTATTTCCA 2399  
Qy 2367 TGGACATTTAGTTAGTGTCTTTTATATACAGGATGATGCTGAGTGACACTCTTGTTGA 2426  
Db 2400 TGGACATTTAGTTAGTGTCTTTTATATACAGGATGATGCTGAGTGACACTCTTGTTGA 2459  
Qy 2427 TATTTCCAAATTTTGTATAGTGTGCTGCACATATTTGAAATCAAAATATTAAGCTTTCC 2486  
Db 2460 TATTTCCAAATTTTGTATAGTGTGCTGCACATATTTGAAATC-ATATATTAAGACTTTCC 2518  
Qy 2487 AAAAATTTGGTCCCTGGTTTTTTCATGGCACTTGATCAGTAAGGATTTCCCTCTGTTTG 2546  
Db 2519 AAAATGAGGTCCCTGGTTTTTTCATGGCACTTGATCAGTAAGGATTTCACTCTGTTTG 2578  
Qy 2547 GAACTAAACCAATTTACTATATGTTAGTAAAGACATTTTTTTTTTCTTCCCTGAAAAA 2606  
Db 2579 TAACTAAACCAATTTACTATATGTTAGTAAAGACATTTCTTTCTCTCTCTCTCTCTCTCT 2638

Qy 2607 -AAAATGAGGAGAGACAAACAAAAA 2636  
Db 2639 TAAAGTGTGGAGAGACAAACAAAAA 2669  
RESULT 4  
ID AAC55715 standard; cDNA; 2669 BP.  
AC AAC55715;  
DT 17-JAN-2001 (first entry)  
XX Human NMB cDNA sequence from Genbank X76534.  
XX Human; differentially regulated gene; macrophage development; diagnosis;  
KW matrix metalloproteinase 19; MMP19; antiarthritic; antiinflammatory;  
KW destructive macrophage development inhibitor; arthritis;  
KW colorectal cancer; immune response; ss.  
XX Homo sapiens.  
XX WO200055373-A2.  
XX 21-SEP-2000.  
XX 15-MAR-2000; 2000WO-US006883.  
XX 15-MAR-1999; 99US-0124530P.  
XX (E0SB-) EOS BIOTECHNOLOGY INC.  
XX Murray R;  
XX WPI; 2000-628200/60.  
Screening drug candidates comprising adding a drug to a cell expressing an  
expression profile gene and determining the effect of the drug on the  
expression of the expression profile gene.  
Claim 1; Page; 99pp; English.  
The present invention describes a method for screening drug candidates.  
The method comprises adding a drug to a cell that expresses an expression  
profile gene encoding a protein encoded by 5 sequences of defined base  
pairs as given in C55638, C55642, C55644 and C55653 or a sequence  
represented by Genbank accession number X92521, X62466, J04130, X62087  
and X76534 (or a fragment) and determining the effect of the drug on the  
expression of the expression profile gene. An inhibitor of matrix  
metalloproteinase 19 (MMP-19), preferably an antibody, is useful for  
treating destructive macrophage disorders (DMD) by inhibiting DM  
development in a cell of an individual having arthritis. Antibodies to  
MMP-19 are useful for localising a therapeutic moiety preferably  
cytotoxic agent or a radioisotope to colorectal cancer tissue. A  
composition comprising MMP-19 is useful for eliciting an immune response  
in an individual. C55635 to C55710 represent human differentially  
regulated genes of the invention. The present sequence represents the  
human NMB cDNA sequence according to the Genbank accession number X76534.  
N.B. The present sequence is not given in the present specification, but  
it is specifically claimed by its Genbank accession number  
Sequence 2669 BP; 752 A; 595 C; 597 G; 725 T; 0 U; 0 Other;

Query Match 95.78; Score 2547.8; DB 3; Length 2669;  
Best Local Similarity 99.28; Pred. No. 0;  
Matches 2591; Conservative 0; Mismatches 17; Indels 3; Gaps 3;  
Qy 28 AACCTTGTGCTGCTCGCTCCGTCGAGAAATTCAGCATGAATGTCTCTACTATTTCCTGGAT 87  
Db 60 ACCTTGAGTGTCTGGTCCGTCGAGAAATTCAGCATGAATGTCTCTACTATTTCCTGGAT 119  
Qy 88 TTCTGCTCTGCTGCTGCAAGATGTCACCTGATGCCCAACGATTTTCATGATGTGCTGG 147

Db 120 TTCTGCTCTGGCTGCAAGATTGCCACTTGATGCCGCAACCAATTTCATGATGCTGTG 179  
QY 148 GCAATGAAGACCTTCTGCTTACATGAGGAGACAAATCAATTAATGAGCTGTCTTCTG 207  
Db 180 GCATGAAGACCTTCTGCTTACATGAGGAGACAAATCAATTAATGAGCTGTCTTCTG 239  
QY 208 ATGAATAAGCTGGAATGAATACTCTACCCAGTGGAGCGGAGAGACATGAGTGA 267  
Db 240 ATGAATAAGCTGGAATGAATACTCTACCCAGTGGAGCGGAGAGACATGAGTGA 299  
QY 268 AAAATCTCTGGAAGGAGCCGTGTGACGCGTCTGACAGTGAATCAACAGCCCTCG 327  
Db 300 AAAATCTCTGGAAGGAGCCGTGTGACGCGTCTGACAGTGAATCAACAGCCCTCG 359  
QY 328 TGGGCTCAAAATATACATTTGGGTGAACCTGATATTCCTAGATGCCAAAAGGAATG 387  
Db 360 TGGGCTCAAAATATACATTTGGGTGAACCTGATATTCCTAGATGCCAAAAGGAATG 419  
QY 388 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATCAGCTGGTTTATCTGCTGATC 447  
Db 420 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 479  
QY 448 CATATGTTTAACTTGGACAGCATGGTTCAGAGGACAGTGCAGCGGAAATGGCACCGGCC 507  
Db 480 CATATGTTTAACTTGGACAGCATGGTTCAGAGGACAGTGCAGCGGAAATGGCACCGGCC 539  
QY 508 AAAGCCATCATAACTCTTCCCTGATGGAAACCTTTTCTCACCACCCCGATGGAGAA 567  
Db 540 AAAGCCATCATAACTCTTCCCTGATGGAAACCTTTTCTCACCACCCCGATGGAGAA 599  
QY 568 GATGAAATTTCACTACGCTCTCCACACACAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 627  
Db 600 GATGAAATTTCACTACGCTCTCCACACACAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 659  
QY 628 GTTCACTGAGAGTTTCTGTGAACACAGCCAAATGTGACACTTGGGCTCAACTCATGGAAG 687  
Db 660 GTTCACTGAGAGTTTCTGTGAACACAGCCAAATGTGACACTTGGGCTCAACTCATGGAAG 719  
QY 688 TGACTGTCTACAGAGACATGACGCGGATATGTTCCCATGCGCAAAAGTGAATGTGT 747  
Db 720 TGACTGTCTACAGAGACATGACGCGGATATGTTCCCATGCGCAAAAGTGAATGTGT 779  
QY 748 ACGTGGTAAACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAGAACGATCGAAAT 807  
Db 780 ACGTGGTAAACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAGAACGATCGAAAT 839  
QY 808 CATCCGACGAAACCTTCC - CAAAGATCTCCCAATTAATGTTGATGTCCTGATTCATGATC 866  
Db 840 CATCCGACGAAACCTTCC - CAAAGATCTCCCAATTAATGTTGATGTCCTGATTCATGATC 899  
QY 867 CTAGCCACTTCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 926  
Db 900 CTAGCCACTTCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 959  
QY 927 GCCTGTTGTTTCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 986  
Db 960 GCCTGTTGTTTCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1019  
QY 987 GCCTTAACCTCACTGTGAAAGCTGACGACGAGGACCTTGTCCGCGCACCGCACCAAC 1046  
Db 1020 GCCTTAACCTCACTGTGAAAGCTGACGACGAGGACCTTGTCCGCGCACCGCACCAAC 1079  
QY 1047 CCAGACCTTCAAAACCCACCCCTTCTTTAGGACCTGTGTTGAGCAACCCCTGAGCTGA 1106  
Db 1080 CCAGACCTTCAAAACCCACCCCTTCTTTAGGACCTGTGTTGAGCAACCCCTGAGCTGA 1139  
QY 1107 GTAGGATTCCTGATGAATACTGCCAGATTAACAGATATGACATTTTCAAGCCACCATCA 1166  
Db 1140 GTAGGATTCCTGATGAATACTGCCAGATTAACAGATATGACATTTTCAAGCCACCATCA 1199  
QY 1167 CAATTGTAGGGGAATCTTAGAGGTTAACTATCATCCAGATGCACAGCTCTGATGCCGG 1226

Db 1200 CAAATTGTAGAGGGAATCTTTAGAGTTTAACTATCATCCAGATGCACAGAGCTCTCTGATGCCG 1259  
QY 1227 TGCCATGGCTGAAAAGCTCCCTAAATAGACTTTTCTGTGACCTGCCAGGAGGATTTCCCA 1286  
Db 1260 TGCCATGGCTGAAAAGCTCCCTAAATAGACTTTTCTGTGACCTGCCAGGAGGATTTCCCA 1319  
QY 1287 CGGAGTCTGTACCATCATATTTCTGACCCCACTGCGAGATCACCCAGAACACAGTCTGCA 1346  
Db 1320 CGGAGTCTGTACCATCATATTTCTGACCCCACTGCGAGATCACCCAGAACACAGTCTGCA 1379  
QY 1347 GCCTGTGTGATGAGATGATGCTCTGCTGACTGTGAGAGCAACCTTCAATGGGTCTG 1406  
Db 1380 GCCTGTGTGATGAGATGATGCTCTGCTGACTGTGAGAGCAACCTTCAATGGGTCTG 1439  
QY 1407 GGAGCTACTGTGTGAACCTTCACTTGGGGATGACACAGCCCTGGCTCTCAGGACACCC 1466  
Db 1440 GGAGCTACTGTGTGAACCTTCACTTGGGGATGACACAGCCCTGGCTCTCAGGACACCC 1499  
QY 1467 TGATTTCTGTTCTGACAGAGACCCAGCTCGCTTTAAGGATGGCAACAGTGCCTGCA 1526  
Db 1500 TGATTTCTGTTCTGACAGAGACCCAGCTCGCTTTAAGGATGGCAACAGTGCCTGCA 1559  
QY 1527 TCTCCGTTGGCTGTGGCCATATTTGTCTGCTGATCTCCCTCTTGGTGTACAAAAAC 1586  
Db 1560 TCTCCGTTGGCTGTGGCCATATTTGTCTGCTGATCTCCCTCTTGGTGTACAAAAAC 1619  
QY 1587 ACAAGGAATAACAACCAATAGAAAATAGTCTGGGAATGTGCTCAGAGCAAGGCTGCA 1646  
Db 1620 ACAAGGAATAACAACCAATAGAAAATAGTCTGGGAATGTGCTCAGAGCAAGGCTGCA 1679  
QY 1647 GTGCTTTCTCAACCGTCAAAAGCCGTGTTCTTCCCGGAAACAGAGAAAGGATCCGC 1706  
Db 1680 GTGCTTTCTCAACCGTCAAAAGCCGTGTTCTTCCCGGAAACAGAGAAAGGATCCGC 1739  
QY 1707 TACTCAAAAACCAAGATTTAAAGAGTTCCTTAAATTTGACCTGTTTCTGAAGCTCA 1766  
Db 1740 TACTCAAAAACCAAGATTTAAAGAGTTCCTTAAATTTGACCTGTTTCTGAAGCTCA 1799  
QY 1767 CTTTTCAGTGCATTTGATGTCAGATGCTCGAGTGGCTATTAACCTTTTCTCTAAAG 1826  
Db 1800 CTTTTCAGTGCATTTGATGTCAGATGCTCGAGTGGCTATTAACCTTTTCTCTAAAG 1859  
QY 1827 ATTAATTGTTAAATAGATTTGCTGTTGGGAAGTGAATTTTTATAGGTTAAATGTCA 1886  
Db 1860 ATTAATTGTTAAATAGATTTGCTGTTGGGAAGTGAATTTTTATAGGTTAAATGTCA 1919  
QY 1887 TTTTAGAGATGGGAGGAGGATTAATCTGACGAGGAGCTTACGCCATGTTGTGAACCTGAT 1946  
Db 1920 TTTTAGAGATGGGAGGAGGATTAATCTGACGAGGAGCTTACGCCATGTTGTGAACCTGAT 1979  
QY 1947 AAAAGCAACTTTAGCAAGCTTCTTTTTCATTTATTTTTATGTTTTCACATTATAAGCTTAG 2006  
Db 1980 AAAAGCAACTTTAGCAAGCTTCTTTTTCATTTATTTTTATGTTTTCACATTATAAGCTTAG 2039  
QY 2007 GTAACGTAGTAGGATAGAAAACACTGTGTCCCGAGAGTAAGAGAGAACTACTATTGATTA 2066  
Db 2040 GTAACGTAGTAGGATAGAAAACACTGTGTCCCGAGAGTAAGAGAGAACTACTATTGATTA 2099  
QY 2067 GAGCTTAACCCAGGTTAACTGCAAGAGAGCGGGATACCTTTCAGCTTTTCAATGTAATG 2126  
Db 2100 GAGCTTAACCCAGGTTAACTGCAAGAGAGCGGGATACCTTTCAGCTTTTCAATGTAATG 2159  
QY 2127 TATGCAATAAGCCAAATGATGCTCAGTTCCTTAAGATCATGTTTCCCAAGCTTAATGATCCCA 2186  
Db 2160 TATGCAATAAGCCAAATGATGCTCAGTTCCTTAAGATCATGTTTCCCAAGCTTAATGATCCCA 2219  
QY 2187 CTTCAATACACATCATGAACTCCTGATGGAAACAATAACAGGCCCAAGCTGTGGTATGA 2246  
Db 2220 CTTCAATACACATCATGAACTCCTGATGGAAACAATAACAGGCCCAAGCTGTGGTATGA 2279  
QY 2247 TGTGCACTTGTAGACTCAGAAAAAATACTACTCTCATAAATGGGTGGGATTTTGT 2306  
Db 2280 TGTGCACTTGTAGACTCAGAAAAAATACTACTCTCATAAATGGGTGGGATTTTGT 2339



Db 900 CTAGCCACTTCTCCTCAATTATTTCTACCAATTAACATAAGTGGAGCTTCGGGGGATAATACTG 959  
Qy 927 GCCTGTTGTTTCCACCATCATCTGTGAATCAGACGCTATGTGCTCAATGAACTTCA 986  
Db 960 GCCTGTTGTTTCCACCATCATCTGTGAATCAGACGCTATGTGCTCAATGAACTTCA 1019  
Qy 987 GCCTTAACCTCACTGTGAAGCTGCAGCACAGGACCTTGTCCGCCACCGCCACCAACAC 1046  
Db 1020 GCCTTAACCTCACTGTGAAGCTGCAGCACAGGACCTTGTCCGCCACCGCCACCAACAC 1079  
Qy 1047 CAGACCTTCAAAACCCACCTTCTTTAGGACCTGTGTCAGCAACCCCTGGAGCTGA 1106  
Db 1080 CAGACCTTCAAAACCCACCTTCTTTAGGACCTGTGTCAGCAACCCCTGGAGCTGA 1139  
Qy 1107 GTAGGATTCCTGATGAAACCTGCCAGATTAACAGATATGCCACCTTCAAGCCACCATCA 1166  
Db 1140 GTAGGATTCCTGATGAAACCTGCCAGATTAACAGATATGCCACCTTCAAGCCACCATCA 1199  
Qy 1167 CAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCCTGATGCCGG 1226  
Db 1200 CAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCCTGATGCCGG 1259  
Qy 1227 TGCCATGGCTGAAGAGTCCCTTAATAGACTTTGTCGTAACCTGCCAAGGGAGCAATCCCA 1286  
Db 1260 TGCCATGGCTGAAGAGTCCCTTAATAGACTTTGTCGTAACCTGCCAAGGGAGCAATCCCA 1319  
Qy 1287 CGGAGTCTGTACCATCATTTCTGACCCCACTGGAGATCACCAGAACACAGTCTGCA 1346  
Db 1320 CGGAGTCTGTACCATCATTTCTGACCCCACTGGAGATCACCAGAACACAGTCTGCA 1379  
Qy 1347 GCCCTGTGGATGTGGATGAGATGTGTCGTGACTGTGAGAGCAACCTTCAATGGTCTG 1406  
Db 1380 GCCCTGTGGATGTGGATGAGATGTGTCGTGACTGTGAGAGCAACCTTCAATGGTCTG 1439  
Qy 1407 GGACGTACTGTGTGAACCTCACCCTGGGGATGACAAAGCCTGCTCAGAGCACCC 1466  
Db 1440 GGACGTACTGTGTGAACCTCACCCTGGGGATGACAAAGCCTGCTCAGAGCACCC 1499  
Qy 1467 TGATTTCTGTTCTTCAGACAGACCCAGCTCGCTTTAAGGATGGCAAAAGTGCCTGA 1526  
Db 1500 TGATTTCTGTTCTTCAGACAGACCCAGCTCGCTTTAAGGATGGCAAAAGTGCCTGA 1559  
Qy 1527 TCTCCGTTGGCTGTTGGCCATATTTGTCAGTGTGATCTCCCTCTGCTGTGACAAAAC 1586  
Db 1560 TCTCCGTTGGCTGTTGGCCATATTTGTCAGTGTGATCTCCCTCTGCTGTGACAAAAC 1619  
Qy 1587 ACAGGAATACAAACCAATAGAAAATAGTCTGGGAATGTGTCAGAAAGCAAGGCCTGA 1646  
Db 1620 ACAGGAATACAAACCAATAGAAAATAGTCTGGGAATGTGTCAGAAAGCAAGGCCTGA 1679  
Qy 1647 GTGCTTCTCAACCGTGCRAAAGCCGTGTTCTTCCCGGGAACACAGAAAGGATCCGC 1706  
Db 1680 GTGCTTCTCAACCGTGCRAAAGCCGTGTTCTTCCCGGGAACACAGAAAGGATCCGC 1739  
Qy 1707 TACTCAAAAACCAAGAAATTTAAAGAGTTCCTTAAATTTGCACTGTTTCTGGAAGCTCA 1766  
Db 1740 TACTCAAAAACCAAGAAATTTAAAGAGTTCCTTAAATTTGCACTGTTTCTGGAAGCTCA 1799  
Qy 1767 CTTTTAGTGCATGATGATGAGATGTGCTGGAGTGCCTAATAACCTTTTTTCTTAAAG 1826  
Db 1800 CTTTTAGTGCATGATGATGAGATGTGCTGGAGTGCCTAATAACCTTTTTTCTTAAAG 1859  
Qy 1827 ATTATCTGTAATAGATATTTGTTGGGAAGTTGAATTTTTTATAGGTTAAATGTC 1886  
Db 1860 ATTATCTGTAATAGATATTTGTTGGGAAGTTGAATTTTTTATAGGTTAAATGTC 1919  
Qy 1887 TTTTAGAGATGGGAGAGGATTTACTGACAGGAGCTTCAGCCATGTTGTGAACCTGAT 1946  
Db 1920 TTTTAGAGATGGGAGAGGATTTACTGACAGGAGCTTCAGCCATGTTGTGAACCTGAT 1979  
Qy 1947 AAAAGCACTTAGCAAGGCTCTTTTCATTATTTTTTATGTTTTCACCTATAAAGCTTAG 2006  
Db 1980 AAAAGCACTTAGCAAGGCTCTTTTCATTATTTTTTATGTTTTCACCTATAAAGCTTAG 2039

Qy 2007 GTAACTAGTAGAATAAACAACCTGTGTCCCGAGAGTAAGAGAGAGCTACTATTGATTA 2066  
Db 2040 GTAACTAGTAGAATAAACAACCTGTGTCCCGAGAGTAAGAGAGAGCTACTATTGATTA 2099  
Qy 2067 GAGCCTAACCCAGGTTAACTGCAAGAGAGGGGGATCTTTCAGCTTCCATGTAACTG 2126  
Db 2100 GAGCCTAACCCAGGTTAACTGCAAGAGAGGGGGATCTTTCAGCTTCCATGTAACTG 2159  
Qy 2127 TATGCATAAAGCAATGTAGTCCAGTTCCTAAGATCATGTTCCCAAGCTAACTGAATCCCA 2186  
Db 2160 TATGCATAAAGCAATGTAGTCCAGTTCCTAAGATCATGTTCCCAAGCTAACTGAATCCCA 2219  
Qy 2187 CTTCAATACACACTCATGAACCTCCTGTATGGAACAATAACAGSCCCTGTGTATGA 2246  
Db 2220 CTTCAATACACACTCATGAACCTCCTGTATGGAACAATAACAGSCCCTGTGTATGA 2279  
Qy 2247 TGTGCACACTTGTAGACTCAGAAAATACTACTCTCATAAATGGGTGGAGTATTTTG 2306  
Db 2280 TGTGCACACTTGTAGACTCAGAAAATACTACTCTCATAAATGGGTGGAGTATTTTG 2339  
Qy 2307 GTGCAACACTTACTTGTGCTGCTGAGTGAAGGAATGATTCATATATTCATTTATTC 2366  
Db 2340 GTGCAACACTTACTTGTGCTGCTGAGTGAAGGAATGATTCATATATTCATTTATTC 2399  
Qy 2367 TGGCACTTACTTGTGCTGCTTATATACAGGAGTATGCTGAGTGACACTCTTGTGTA 2426  
Db 2400 TGGCACTTACTTGTGCTGCTTATATACAGGAGTATGCTGAGTGACACTCTTGTGTA 2459  
Qy 2427 TATTTCCAAATTTTGTATAGTCCGTCGACATATTTGAAATCAAAATTAAGACTTTCC 2486  
Db 2460 TATTTCCAAATTTTGTATAGTCCGTCGACATATTTGAAATCAAAATTAAGACTTTCC 2518  
Qy 2487 AAAAATTTGTCCTGCTGCTTTCATGCAACTGATCAGTAAGGATTTCCCTCTGTTTG 2546  
Db 2519 AAAGATGAGTCCCTGCTTTCATGCAACTGATCAGTAAGGATTTCCCTCTGTTTG 2578  
Qy 2547 GAACTAAACCATTTACTATATAGTTAGACAGACATTTTTTTTTTCTTCTCTGAAAAA 2606  
Db 2579 TAACTAAACCATCTACTATATGTTAGACATGACATTTCTTCTCTCTCTCTGAAAAA 2638  
Qy 2607 -AAATGAGGAGAGACAAAAA 2636  
Db 2639 TAAAGTGGGAGAGACAAAAA 2669

## RESULT 6

ABX76321  
ID ABX76321 standard; DNA; 2669 BP.

XX AC ABX76321;

XX DT 02-APR-2003 (first entry)

XX DE Lung cancer-associated polynucleotide #185.

XX KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema;  
XX KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;  
XX KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
XX KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
XX KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX OS Unidentified.

XX PN WO200286443-A2.

XX PD 31-OCT-2002.

XX PF 18-APR-2002; 2002WO-US012476.

XX PR 18-APR-2001; 2001US-0284770P.

XX PR 10-MAY-2001; 2001US-0290492P.

XX PR 09-NOV-2001; 2001US-0339245P.







Db 1560 TCTCCGTTGGCTGCTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTCACAAAAC 1619  
QY 1587 ACAAGGAATACACCAACCAATAGAAATAGTCTCTGGGAATGTGTGTCAGAACCAAGGCTGA 1646  
Db 1620 ACAGGAATACACCAACCAATAGAAATAGTCTCTGGGAATGTGTGTCAGAACCAAGGCTGA 1679  
QY 1647 GTGTCTTTCTCAACCGTCAAAAGCCGTGTCTTCCCGGGAACACAGGAAGGATCCGC 1706  
Db 1680 GTGTCTTTCTCAACCGTCAAAAGCCGTGTCTTCCCGGGAACACAGGAAGGATCCGC 1739  
QY 1707 TACTCAAAAACCAAGAAATTAAGAGGTTCTTAAATTCGACCTTGTTCTCAAGCTCA 1766  
Db 1740 TACTCAAAAACCAAGAAATTAAGAGGTTCTTAAATTCGACCTTGTTCTCAAGCTCA 1799  
QY 1767 CTTTTCAGTGGCATTGATGTGAGATGCTGTGAGTGGCTATTAACTTTTCTTAAG 1826  
Db 1800 CTTTTCAGTGGCATTGATGTGAGATGCTGTGAGTGGCTATTAACTTTTCTTAAG 1859  
QY 1827 ATTATTGTTAAATAGATATTGTGGTTTGGGGAAGTTGAATTTTATAGTTAAATGTCA 1886  
Db 1860 ATTATTGTTAAATAGATATTGTGGTTTGGGGAAGTTGAATTTTATAGTTAAATGTCA 1919  
QY 1887 TTTTAGATGGGAGAGGATTAATCTGACGCGAGCTTCACCATGTTGTCAAACTGAT 1946  
Db 1920 TTTTAGATGGGAGAGGATTAATCTGACGCGAGCTTCACCATGTTGTCAAACTGAT 1979  
QY 1947 AAAAGCAACTAGCAAGGCTCTTTTCAATTTTATGTTTCACTTATAAAGTCTTAG 2006  
Db 1980 AAAAGCAACTAGCAAGGCTCTTTTCAATTTTATGTTTCACTTATAAAGTCTTAG 2039  
QY 2007 GTAACTAGTAGATAGAAACACTGTGTCCGAGAGTAAGGAGAGACTACTATTGATTA 2066  
Db 2040 GTAACTAGTAGATAGAAACACTGTGTCCGAGAGTAAGGAGAGACTACTATTGATTA 2099  
QY 2067 GAGCCTAACCCAGGTTAACTGCAAGAGAGCGGATACCTTTTCCATGTAACGTG 2126  
Db 2100 GAGCCTAACCCAGGTTAACTGCAAGAGAGCGGATACCTTTTCCATGTAACGTG 2159  
QY 2127 TATGATAAAGCCATGTAGTCCAGTTTCTAAGATCATGTTCCAAAGCTGAATCCCA 2186  
Db 2160 TATGATAAAGCCATGTAGTCCAGTTTCTAAGATCATGTTCCAAAGCTGAATCCCA 2219  
QY 2187 CTTCAATACACACTCATGAATCTCTGTGGAACAAATACAGGCCCAAGCCCTGTGTATGA 2246  
Db 2220 CTTCAATACACACTCATGAATCTCTGTGGAACAAATACAGGCCCAAGCCCTGTGTATGA 2279  
QY 2247 TGTGCACTTGTAGTCTAGAAATACTCTCTATAAATGGGTGGGAGTATTTTG 2306  
Db 2280 TGTGCACTTGTAGTCTAGAAATACTCTCTATAAATGGGTGGGAGTATTTTG 2339  
QY 2307 GTGCAACTTGTCTTGTGCTGAGTGAAGGAATGATTTTCAATATTCATTATTCCTCA 2366  
Db 2340 GTGCAACTTGTCTTGTGCTGAGTGAAGGAATGATTTTCAATATTCATTATTCCTCA 2399  
QY 2367 TGGCACTTGTAGTCTTGTATATATACAGGATGATGCTGAGTGACACTCTTGTGTA 2426  
Db 2400 TGGCACTTGTAGTCTTGTATATACAGGATGATGCTGAGTGACACTCTTGTGTA 2459  
QY 2427 TATTTCAAAATTTTGTATAGTCTGCTGCACATATTTGGAATCAAAATATTAAGACTTTCC 2486  
Db 2460 TATTTCAAAATTTTGTATAGTCTGCTGCACATATTTGGAATC-ATATATTAAGACTTTCC 2518  
QY 2487 AAAAATTTGGTCCCTGTTTTCATGGAACCTTGAATCAGTAAGATTTCCCTCTGTTTG 2546  
Db 2519 AAAGATGAGGTCCTGTTTTCATGGAACCTTGAATCAGTAAGATTTCACTCTGTTTG 2578  
QY 2547 GAATCAAAACCACTTACTATATGTTAGACAGACATTTTTTTTTTTTCTTCTCGTGAATA 2606  
Db 2579 TAATCAAAACCACTTACTATATGTTAGACAGACATTTTTTTTCTTCTCGTGAATA 2638  
QY 2607 -AAATGAGGAAGAGACAAAAAATAAAAAA 2636  
Db 2639 TAAAGTGTGGAGAGACAAAAAATAAAAAA 2669

RESULT 7  
ADD78274

ID ADD78274 standard; DNA; 2728 BP.

XX ADD78274;

AC ADD78274;

XX 29-JAN-2004 (first entry)

XX Human CGDD-16 coding sequence.

XX Anabolic; Hypertensive; Respiratory; Anti-HIV; Antiallergic;  
KW Neuroprotective; Nootropic; Antianemic; Antiarteriosclerotic;  
KW Antiinflammatory; Ophthalmological; Muscular; Hepatotropic;  
KW Neuroprotective; Antiaesthetic; Anticonvulsant; Virucide; Antibacterial;  
KW Fungicide; Antiparasitic; Protozoacide; Antihelminthic; Cytostatic;  
KW Cerebroprotective; Antiparkinsonian; Antipsoriatic; Anticout;  
KW Antidiabetic; Antiarthritic; Antirheumatic; Osteopathic; Gene therapy;  
KW human; cell growth; cell differentiation; cell death; CGDD;  
KW cell proliferative disorder; cancer; developmental disorder;  
KW neurological disorder; autoimmune disorder; inflammatory disorder;  
KW infection; reproductive disorder; gene; ds.

XX Homo sapiens.

XX WO2003077875-A2.

XX 25-SEP-2003.

XX 14-MAR-2003; 2003WO-US008310.

XX 15-MAR-2002; 2002US-0364494P.

XX 29-MAR-2002; 2002US-0369129P.

XX 12-APR-2002; 2002US-0372511P.

XX (INCY-) INCYTE GENOMICS INC.

XX Kable AE, Tran UK, Hafalia AJA, Burford N, Honchell CD;  
PI Lehr-Mason PM, Duggan BM, Ramkumar J, Griffin JA, Richardson TW;  
PI Elliott VS, Jiang X, Jackson AA, Marquis JP, Chawla NK, Khare R;  
PI Becha SD, Lee ST, Swarnakar A, Yue H, Warren BA, Baughn MR, Lal PG;  
PI Lee S, Ho A, Gandhi AR, Yao MG;

XX WPI; 2003-779081/73.

XX P-PSDB; ADD78235.

XX New polypeptides and polynucleotides associated with cell growth,  
PT differentiation and death, useful for diagnosing, treating or preventing  
PT e.g. developmental, neurological, autoimmune, inflammatory or  
PT reproductive disorders.

XX Claim 5; SEQ ID NO 55; 320pp; English.

XX The present invention relates to novel human proteins (I; ADD78220-  
CC ADD78258) and their coding sequences (II; ADD78259-ADD78297), which are  
CC associated with cell growth, differentiation and death, referred to as  
CC CGDD-n proteins, where n is a number from 1 to 39. The CGDD proteins and  
CC their coding sequences are useful for diagnosing, treating or preventing  
CC cell proliferative disorders (e.g. cirrhosis, hepatitis,  
CC arteriosclerosis, psoriasis, primary thrombocytopenia) or cancers (e.g.  
CC adenocarcinoma, sarcoma or cancers of the bone, bone marrow, brain,  
CC breast, colon, kidney, liver, lung or uterus), developmental disorders  
CC (e.g. renal tubular acidosis, Becker muscular dystrophy, gonadal  
CC dysgenesis, hypothyroidism or seizures), neurological disorders (e.g.  
CC Pick's disease, cataplexy, epilepsy, ischemic cerebrovascular disease,  
CC stroke, Alzheimer's disease, Parkinson's disease or dementia),  
CC autoimmune/inflammatory disorders (e.g. AIDS, allergies, anemia, asthma,  
CC diabetes mellitus, bronchitis, osteoporosis, osteoarthritis, rheumatoid  
CC arthritis, contact dermatitis or gout), viral, bacterial, fungal,  
CC parasitic, protozoan or helminthic infections, reproductive disorders  
CC (e.g. infertility, ectopic pregnancy, premature ovarian failure, delayed  
CC puberty or prostatitis) or disorders of the placenta (e.g. preeclampsia,

CC	choriocarcinoma, placenta previa, placental or maternal floor infarction	
CC	or chronic villitis).	
XX		
SQ	Sequence 2728 BP; 778 A; 610 C; 607 G; 733 T; 0 U; 0 Other;	
	Query Match 94.2%; Score 2505.6; DB 9; Length 2728;	
	Best Local Similarity 97.8%; Pred. No. 0;	
	Matches 2596; Conservative 0; Mismatches 19; Indels 39; Gaps 4;	
QY	28 AACCTTGGTGGCTGGTGGTGGAGAAATTCAGCATGGAATGTCTCTACTATTTCCTGGGAT 87	987 GCCTTAACCTCACTGTGTAAGAGCTGCAGCACAGAGACCTTGTCCGCCACCGCCACCACAC 1046
DB	76 ACCTTGGTGGCTGGTGGTGGAGAAATTCAGCATGGAATGTCTCTACTATTTCCTGGGAT 135	1036 GCCTTAACCTCACTGTGTAAGAGCTGCAGCACAGAGACCTTGTCCGCCACCGCCACCACAC 1095
QY	88 TTCTGCTCTGCTGCTGCAAGATTGCCATTGTATGCCGCAAGATTTCATGATGTCTGG 147	1047 CCAGACCTTGAAGACCCACCCCTTCTTT----- 1074
DB	136 TTCTGCTCTGCTGCTGCAAGATTGCCATTGTATGCCGCAAGATTTCATGATGTCTGG 195	1096 CCAGACCTTGAAGACCCACCCCTTCTTTAGCAACTACTCTMAAATCTTATGATTTCAACA 1155
QY	148 GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAATCAATTAATGCTGCTTTCTG 207	1075 ----AGGACCTGCTGGTGACAAACCCCTCGAGCTCAGTAGGATTCCTGATGAAAATCTGCC 1130
DB	196 GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAATCAATTAATGCTGCTTTCTG 255	1156 CCCAGAGACCTGCTGGTGACAAACCCCTCGAGCTCAGTAGGATTCCTGATGAAAATCTGCC 1215
QY	208 ATGAAATGACTGGAATGAAACTCTACCCAGTGTGGAGGGGAGACATGAGTGGA 267	1131 AGATTAA CAGATATGGCCACTTTTAAAGCCACCATCAAAATTTGTAGAGGAAATCTTAGAG 1190
DB	256 ATGAAATGACTGGAATGAAACTCTACCCAGTGTGGAGGGGAGACATGAGTGGA 315	1216 AGATTAA CAGATATGGCCACTTTTAAAGCCACCATCAAAATTTGTAGAGGAAATCTTAGAG 1275
QY	268 AAAACTCTCGGAAGGAGCGGTGTGACGCGGTCTGACCACTGACTCACCAGCCCTCG 327	1191 TTACATCATCCAGATGACAGACGCTCTGATGCCGGTGCATGGCTGAAAGCTCCCTAA 1250
DB	316 AAAACTCTCGGAAGGAGCGGTGTGACGCGGTCTGACCACTGACTCACCAGCCCTCG 375	1276 TTACATCATCCAGATGACAGACGCTCTGATGCCGGTGCATGGCTGAAAGCTCCCTAA 1335
QY	328 TGGGCTCAAAATATAACATTTGCGGTGAACCTGATATTCCTAGATGCCAAAAGGAGATG 387	1251 TAGACTTTGTCTGTGACCTGCCAAGGAGCATTTCCACGAGGTCTGTACCATCATTTCTG 1310
DB	376 TGGGCTCAAAATATAACATTTGCGGTGAACCTGATATTCCTAGATGCCAAAAGGAGATG 435	1336 TAGACTTTGTCTGTGACCTGCCAAGGAGCATTTCCACGAGGTCTGTACCATCATTTCTG 1395
QY	388 CCAATGGCAACATAGTCTATGAGAGAACTGAGAAATGAGGTGTTTATCTGCTGATC 447	1311 ACCCACCTGCGAGATCACCAAGAACACAGTCTGCAGCCCTGTGGATGTGGATGATGT 1370
DB	436 CCAATGGCAACATAGTCTATGAGAGAACTGAGAAATGAGGTGTTTATCTGCTGATC 495	1396 ACCCACCTGCGAGATCACCAAGAACACAGTCTGCAGCCCTGTGGATGTGGATGATGT 1455
QY	448 CATATGTTTACACTGGACAGCATGTGACAGGACAGTACGCGGGGAAAATGGCACCGGC 507	1371 GTCTGCTGACTGTGAGACGAACTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCC 1430
DB	496 CGTATGTTTACACTGGACAGCATGTGACAGGACAGTACGCGGGGAAAATGGCACCGGC 555	1456 GTCTGCTGACTGTGAGACGAACTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCC 1515
QY	508 AAAGCCATCATACAGTCTTCCCTGATGGGAAACCTTTTCTCACCACCCCGGATGAGAA 567	1431 TGGGGGATGACACAAAGCCTGGCTCTCAGAGACACCTGTCTGTCTGTCTGACAGAGACC 1490
DB	556 AAAGCCATCATACAGTCTTCCCTGATGGGAAACCTTTTCTCACCACCCCGGATGAGAA 615	1516 TGGGGGATGACACAAAGCCTGGCTCTCAGAGACACCTGTCTGTCTGTCTGACAGAGACC 1575
QY	568 GATGGAATTTCACTACGCTTCCACACTTGGTCACTATTTCCAGAAATTTGGACGAT 627	1491 CAGCTTCGCTTTAAGGATGGCAAAACAGTGCCTGTATCTCCGTTCGCTGCTGGCCATAT 1550
DB	616 GATGGAATTTCACTACGCTTCCACACTTGGTCACTATTTCCAGAAATTTGGACGAT 675	1576 CAGCTTCGCTTTAAGGATGGCAAAACAGTGCCTGTATCTCCGTTCGCTGCTGGCCATAT 1635
QY	628 GTTCAGTGAGATTTCTGTGAACACAGCCAAATGTGACACTTGGGCTCAACTCATGGAAG 687	1551 TTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACACAAAGGAATACACCCCAATAGAA 1610
DB	676 GTTCAGTGAGATTTCTGTGAACACAGCCAAATGTGACACTTGGGCTCAACTCATGGAAG 735	1636 TTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACACAAAGGAATACACCCCAATAGAA 1695
QY	688 TGACTGTCTACAGAGACATGACGGGCATATGTTCCCATCGCAAGTGAAGATGTGT 747	1611 ATAGTCTCGGGAATGTGTTCAGAAGCAAAAGGCTGAGTGTCTTTCTCAACCGTGCAAAAG 1670
DB	736 TGACTGTCTACAGAGACATGACGGGCATATGTTCCCATCGCAAGTGAAGATGTGT 795	1696 ATAGTCTCGGGAATGTGTTCAGAAGCAAAAGGCTGAGTGTCTTTCTCAACCGTGCAAAAG 1755
QY	748 ACCTGGTAACAGATCAGATTCCTGTTGTGACTATGTTCCAGAGAACGATCGAAAT 807	1671 CCGTGTCTTCCCGGAAACACAGGAAAGGATCCGCTACTCRAAAACCAAGAAATTTAAAG 1730
DB	796 ACCTGGTAACAGATCAGATTCCTGTTGTGACTATGTTCCAGAGAACGATCGAAAT 855	1756 CCGTGTCTTCCCGGAAACACAGGAAAGGATCCGCTACTCRAAAACCAAGAAATTTAAAG 1815
QY	808 CATCCGACGAAACCTTCC - CAAAGATCTCCCAATATGTTGTGATGCTCCTGATTCATGATC 866	1731 GAGTTCCTTAAATTTTCGACCTGTTTCTGAAGCTCACTTTTCAGTGCCATTTGATGTGAGA 1790
DB	856 CATCCGACGAAACCTTCCCAAGATCTCCCAATATGTTGTGATGCTCCTGATTCATGATC 915	1816 GAGTTCCTTAAATTTTCGACCTGTTTCTGAAGCTCACTTTTCAGTGCCATTTGATGTGAGA 1875
QY	867 CTAGCCACTTCTCAATATTTCTCAATTAACATCAAGTGAGCTTCGGGATAAATCTG 926	1791 TGTGCTGGAGTGGCTATTAACTTTTTCCTAAAGATTAATGTTTAAATAGATTAATGTGG 1850
DB	916 CTAGCCACTTCTCAATATTTCTCAATTAACATCAAGTGAGCTTCGGGATAAATCTG 975	1876 TGTGCTGGAGTGGCTATTAACTTTTTCCTAAAGATTAATGTTTAAATAGATTAATGTGG 1935
QY	927 GCCTGTTTGTTCACCAATCATACTGTAATCACAGTATGCTCAATGGACCTTCA 986	1851 TTTGGGGAAGTGGAAATTTTATAGGTTAAATGTCATTTTAGAGATGGGAGGGGATTA 1910
DB	976 GCCTGTTTGTTCACCAATCATACTGTAATCACAGTATGCTCAATGGACCTTCA 1035	1936 TTTGGGGAAGTGGAAATTTTATAGGTTAAATGTCATTTTAGAGATGGGAGGGGATTA 1995

```
Db 2116 TGTCCGAGAGTAAGGAGAGAGCTACTATTGATTAGAGCCTAACCCAGGTTAACTGCAA 2175
Qy 2091 GAAGAGGGGGATATCTTTCAGCTTTCCATGTAACCTGTATGCAATAAGCAATGTAGTCCA 2150
Db 2176 GAAGAGGGGGATATCTTTCAGCTTTCCATGTAACCTGTATGCAATAAGCAATGTAGTCCA 2235
Qy 2151 GTTCTTAAGATCATGTTCCAGAGTAACTGAATCCCACTTCAATACACACTCATGAACCTCC 2210
Db 2236 GTTCTTAAGATCATGTTCCAGAGTAACTGAATCCCACTTCAATACACACTCATGAACCTCC 2295
Qy 2211 TGATGGAAACAATAACAGGCCCAAGCCTGTGTATGATGTCACACTTGTAGACTCAGAA 2270
Db 2296 TGATGGAAACAATAACAGGCCCAAGCCTGTGTATGATGTCACACTTGTAGACTCAGAA 2355
Qy 2271 AAAATACTACTCTCTATAAATGGTGGGAGTATTTGGTGACAACCTATTTGGCTGGCTG 2330
Db 2356 AAAATACTACTCTCTATAAATGGTGGGAGTATTTGGTGACAACCTATTTGGCTGGCTG 2415
Qy 2331 AGTGAAGGAATGATATTCATATATTCATTTATTCATGACATTTAGTAGTGTCTTTTA 2390
Db 2416 AGTGAAGGAATGATATTCATATATTCATTTATTCATGACATTTAGTAGTGTCTTTTA 2475
Qy 2391 TATACAGGCAATGATGCTGAGTGACACTCTTGTGTATATTTCCAAATTTTGTATAGTCG 2450
Db 2476 TATACAGGCAATGATGCTGAGTGACACTCTTGTGTATATTTCCAAATTTTGTATAGTCG 2535
Qy 2451 CTGCACATATTTGAATCAAAATATTAAGACTTTCACAAAATTTGGTCCCTGGTTTTTCA 2510
Db 2536 CTGCACATATTTGAATC-ATATATTAAGACTTTCACAAAATTTGGTCCCTGGTTTTTCA 2594
Qy 2511 TGGCAACTTGATCAGTAGGATTTCCCTCTGTTTGGAACTAAACCAATTTACTATATGT 2570
Db 2595 TGGCAACTTGATCAGTAGGATTTCCCTCTGTTTGGAACTAAACCAATTTACTATATGT 2654
Qy 2571 TAGCAAGACATTTTTTTTTTTTCTCTCTGAAAAA-AAAAATGAGGGAGAGACAAAAA 2629
Db 2655 TAGCAATGACATTTCTTTCTCTCTCTCTGAAAAATTAAGTGGGAGAGACAAAAA 2714
Qy 2630 AAAAAAAAAAAAAA 2643
Db 2715 AAAAAAAAAAAAAA 2728
```

## RESULT 8

ABK33556

ID ABK33556 standard; cDNA; 2683 BP.

XX AC ABK33556;

XX XX 08-MAY-2002 (first entry)

XX DE cDNA encoding human PRO protein, Seq ID No 41.

XX KW Human; secreted protein; PRO; tumour; lung cancer; colon cancer;  
XX KW breast cancer; prostate tumour; rectal tumour; liver tumour;  
XX KW pericyte cell proliferation; chondrocyte cell proliferation;  
XX KW tumour necrosis factor-alpha; gene; ss.

XX OS Homo sapiens.

XX XX W0200208288-A2.

XX XX 31-JAN-2002.

XX XX 29-JUN-2001; 2001WO-US021066.

XX XX 20-JUL-2000; 2000US-0219556P.

XX XX 25-JUL-2000; 2000US-0220585P.

XX XX 25-JUL-2000; 2000US-0220605P.

XX XX 25-JUL-2000; 2000US-0220607P.

XX XX 25-JUL-2000; 2000US-0220624P.

XX XX 25-JUL-2000; 2000US-0220638P.

PR 25-JUL-2000; 2000US-0220664P.  
PR 25-JUL-2000; 2000US-0220666P.  
PR 26-JUL-2000; 2000US-0220893P.  
PR 28-JUL-2000; 2000WO-US020710.  
PR 01-AUG-2000; 2000US-0222425P.  
PR 22-AUG-2000; 2000US-0227133P.  
PR 23-AUG-2000; 2000WO-US023522.  
PR 24-AUG-2000; 2000WO-US023328.  
PR 10-NOV-2000; 2000WO-US030873.  
PR 28-NOV-2000; 2000US-0253646P.  
PR 01-DEC-2000; 2000WO-US032678.  
PR 20-DEC-2000; 2000US-00747259.  
PR 20-DEC-2000; 2000WO-US034956.  
PR 28-FEB-2001; 2001WO-US006520.  
PR 01-MAR-2001; 2001WO-US006666.  
PR 22-MAR-2001; 2001US-00816744.  
PR 10-MAY-2001; 2001US-00854208.  
PR 10-MAY-2001; 2001US-00854280.  
PR 25-MAY-2001; 2001WO-US017092.

(GETH ) GENENTECH INC.

Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ,  
Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

WPI; 2002-172001/22.

P-PSDB; AAU83612.

One hundred and twenty two nucleic acids encoding PRO polypeptides,  
useful for treating a PRO related disorder and for diagnosing tumors such  
as lung cancer, colon cancer, breast tumor, prostate tumor, rectal tumor  
or liver tumor.

Claim 2; Fig 41; 359pp; English.

The invention relates to one hundred and twenty two nucleic acids  
encoding PRO polypeptides. The sequences of the 122 PRO polynucleotides  
encode human secreted proteins. The PRO nucleic acids, polypeptides,  
agonists and antagonists are useful for treating a PRO related disorder.  
The PRO polypeptides are useful for diagnosing tumours, especially lung  
cancer, colon cancer, breast tumour, prostate tumour, rectal tumour or  
liver tumour. The PRO polypeptides are useful for stimulating the  
proliferation of, or gene expression, in pericyte cells, for stimulating  
the proliferation or differentiation of chondrocyte cells, for  
stimulating the release of tumour necrosis factor-alpha from human blood,  
for stimulating or inhibiting the proliferation of normal human dermal  
fibroblast cells. The PRO polypeptide may also be used as molecular  
weight markers and for tissue typing. The PRO nucleic acids have  
applications in molecular biology, including use as hybridisation probes,  
and in chromosome and gene mapping. ABK33536-ABK33657 represent human PRO  
protein coding sequences of the invention

SQ Sequence 2683 BP; 751 A; 604 C; 596 G; 732 T; 0 U; 0 Other;

Query Match

Best Local Similarity 93.4%; Score 2485; DB 6; Length 2683;

Matches 2576; Conservative 0; Mismatches 20; Indels 39; Gaps 4;

Qy 28 AACCTGGTGCCTGCGTCGTGAGAAATTCAGCATGGAATCTCTACTATTCTCGGAT 87

Db 50 ACCTTGAGTGCCTGCGTCGTGAGAAATTCAGCATGGAATCTCTACTATTCTCGGAT 109

Qy 88 TTCTGCTCTCGGCTGCAAGATTGCCACTTGTATCCGCCCAACGATTTCATGATGCTGG 147

Db 110 TTCTGCTCTCGGCTGCAAGATTGCCACTTGTATCCGCCCAACGATTTCATGATGCTGG 169

Qy 148 GCAATGAAGACCTTCTGCTTACATGAGGAGACAAATCAATTAATGGCTGCTTCTG 207

Db 170 GCAATGAAGACCTTCTGCTTACATGAGGAGACAAATCAATTAATGGCTGCTTCTG 229

Qy 208 ATGAAATGACTGGAATGAAAATCTACCCAGTGTGGAAGCGGGAGACATGAGTGA 267

Db 230 ATGAAATGACTGGAATGAAAATCTACCCAGTGTGGAAGCGGGAGACATGAGTGA 289

QY	268	AAAACTCCTGGNAGGGAGCCGTGTGCGAGGGGTCCTGACAGTGACTCACCAGCCCTCG	327	1311	ACCCACCTCGGAGATCACCAGAACACAGTCTCGAGCCCTGTGGATGTGGATGAGATGT	1370
Db	290	AAAACTCCTGGNAGGGAGCCGTGTGCGAGGGGTCCTGACAGTGACTCACCAGCCCTCG	349	1370	ACCCACCTCGGAGATCACCAGAACACAGTCTCGAGCCCTGTGGATGTGGATGAGATGT	1429
QY	328	TGGGCTCAAAATATAACATTTGGGGTGAACCTGATATTCCTTAGATGCCAAAAGGAAGATG	387	1371	GTCTGCTGACTGTGAGAGCAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCC	1430
Db	350	TGGGCTCAAAATATAACATTTGGGGTGAACCTGATATTCCTTAGATGCCAAAAGGAAGATG	409	1430	GTCTGCTGACTGTGAGAGCAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCC	1489
QY	388	CCAAATGGCAACATAGTCTATGAGAGAACTGCGAGAAATGAGGCTGGTTATCTGCTGATC	447	1431	TGGGGGATGACACAAAGCCTGCTCTCACGAGCACCTGATTTCTCTCTCTGACAGAGACC	1490
Db	410	CCAAATGGCAACATAGTCTATGAGAGAACTGCGAGAAATGAGGCTGGTTATCTGCTGATC	469	1490	TGGGGGATGACACAAAGCCTGCTCTCACGAGCACCTGATTTCTCTCTCTGACAGAGACC	1549
QY	448	CATATGTTTACAACTGGACAGCATGCTCAGAGGACAGTGCAGGGGAAAATGSCACCGGCC	507	1491	CAGCCTCGCCTTTTAAGGATGGCAACACAGTGCCTGATCTCCGTTGGCTGTCTGGCCATAT	1550
Db	470	CGTATGTTTACAACTGGACAGCATGCTCAGAGGACAGTGCAGGGGAAAATGSCACCGGCC	529	1550	CAGCCTCGCCTTTTAAGGATGGCAACACAGTGCCTGATCTCCGTTGGCTGTCTGGCCATAT	1609
QY	508	AAAGCCATCATAACTCTTCCCTGATGGGAAACCTTTTCTCACCACCCCGGATGGAGAA	567	1551	TTGTCACTGTGATCTCCCTCTTGGTGTACAAAACACAAAGGAATACAAACCAATAGAAA	1610
Db	530	AAAGCCATCATAACTCTTCCCTGATGGGAAACCTTTTCTCACCACCCCGGATGGAGAA	589	1610	TTGTCACTGTGATCTCCCTCTTGGTGTACAAAACACAAAGGAATACAAACCAATAGAAA	1669
QY	568	GATGAAATTTCACTACGCTTCCACACACTTTGGTCAAGTATTTCCAGAAATTTGGGACGAT	627	1611	ATAGTCTCTGGGAAATGTGGTCAAGCAAAAGGCTGAGTGTCTTTCTCAAACCGTCAAAAG	1670
Db	590	GATGAAATTTCACTACGCTTCCACACACTTTGGTCAAGTATTTCCAGAAATTTGGGACGAT	649	1670	ATAGTCTCTGGGAAATGTGGTCAAGCAAAAGGCTGAGTGTCTTTCTCAAACCGTCAAAAG	1729
QY	628	GTTCAGTGAGAGTTTCTGTGAAACAGCCCAATGTGACACTTGGGCTCAACTCATGGGAAG	687	1671	CCGTGTTCTTCCCGGAAAACAGGAAAGGATCCGCTACTCAAAAACCAAGAAATTTAAAG	1730
Db	650	GTTCAGTGAGAGTTTCTGTGAAACAGCCCAATGTGACACTTGGGCTCAACTCATGGGAAG	709	1730	CCGTGTTCTTCCCGGAAAACAGGAAAGGATCCGCTACTCAAAAACCAAGAAATTTAAAG	1789
QY	688	TCGACTGTCTACAGAAACATGAGCGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT	747	1731	GAGTTTCTTAATTTTCGACCTTGTCTGAACTCACATTTTCAGTGCCATTCATGTGAGA	1790
Db	710	TCGACTGTCTACAGAAACATGAGCGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT	769	1790	GAGTTTCTTAATTTTCGACCTTGTCTGAACTCACATTTTCAGTGCCATTCATGTGAGA	1849
QY	748	ACGTGTAACAGATCAGATTCCTGTGTTTGTGACTATGTTCCAGAAAGCATCGAAATTT	807	1791	TCGTCTGGAGTGGCTATTAACCTTTTTCCTTAAAGATTAATTTGTTAAATAGATATGTGG	1850
Db	770	ACGTGTAACAGATCAGATTCCTGTGTTTGTGACTATGTTCCAGAAAGCATCGAAATTT	829	1850	TCGTCTGGAGTGGCTATTAACCTTTTTCCTTAAAGATTAATTTGTTAAATAGATATGTGG	1909
QY	808	CATCCGACAAACCTTCC-CAAGATCTCCCATATGTTTGATGTCCTGATTCATGATC	866	1851	TTTGGGGAAGTTGAAATTTTATAGGTTAAATGTCAATTTTAGAGATGGGGAGGGATTA	1910
Db	830	CATCCGACAAACCTTCC-CAAGATCTCCCATATGTTTGATGTCCTGATTCATGATC	889	1910	TTTGGGGAAGTTGAAATTTTATAGGTTAAATGTCAATTTTAGAGATGGGGAGGGATTA	1969
QY	867	CTAGCCACTTCTCAATATTCTACCAATTAACCTACAAAGTGGAGCTTCGGGGATTAATCTG	926	1911	TTACTGAGCAGCTTCAGCCATGTTGTGAACTGATTAAGCAACTTACAGAGGCTTCTT	1970
Db	890	CTAGCCACTTCTCAATATTCTACCAATTAACCTACAAAGTGGAGCTTCGGGGATTAATCTG	949	1970	TACTGAGCAGCTTCAGCCATGTTGTGAACTGATTAAGCAACTTACAGAGGCTTCTT	2029
QY	927	GCTGTTTGTTCACCAATCATCTGTGAATCACAAGTATGTGCTCAATGGAACCTTCA	986	1971	TTTCAATTTTATTTATGTTTCACTTATAAGTCTTAGGTAACCTAGTAGGATAGAAAACCTG	2030
Db	950	GCTGTTTGTTCACCAATCATCTGTGAATCACAAGTATGTGCTCAATGGAACCTTCA	1009	2030	TTTCAATTTTATTTATGTTTCACTTATAAGTCTTAGGTAACCTAGTAGGATAGAAAACCTG	2089
QY	987	GCTTAACTCTACTGTGAAGCTGCAGCACAGGACCTTGTCCGCCACCGCCACCCAC	1046	2031	TGTCCCGAGAGTAAGGAGAGAAAGCTACTATTGATTTAGAGCCTAAACCCAGGTTAACTGCAA	2090
Db	1010	GCTTAACTCTACTGTGAAGCTGCAGCACAGGACCTTGTCCGCCACCGCCACCCAC	1069	2090	TGTCCCGAGAGTAAGGAGAGAAAGCTACTATTGATTTAGAGCCTAAACCCAGGTTAACTGCAA	2149
QY	1047	CCAGACCTTCAAAACCCACCCCTTCTTT-----	1074	2091	GAAGAGCGGGATACCTTCAGCTTTCCATGTAACCTGATGATAAGGCAATGATGTGCA	2150
Db	1070	CCAGACCTTCAAAACCCACCCCTTCTTTTAGCAACTACTCTAAATCTTATGATTCAAAACA	1129	2150	GAAGAGCGGGATACCTTCAGCTTTCCATGTAACCTGATGATAAGGCAATGATGTGCA	2209
QY	1075	----AGACCTGCTGGTGACAAACCCCTGGAGCTGATGAGGATTCCTGATGAAACTGCTGC	1130	2151	GTTTCTTAAGATCATGTTTCCAAAGCTAACTGAATCCCACTTCAATACACACTCATGAACCTC	2210
Db	1130	CCCGAGGACCTACTGGTGAACACCCCTGGAGCTGATGAGGATTCCTGATGAAACTGCTGC	1189	2210	GTTTCTTAAGATCATGTTTCCAAAGCTAACTGAATCCCACTTCAATACACACTCATGAACCTC	2269
QY	1131	AGATTAAACAGATATGGCCATTTTCAGCCACCATCAATTTGTAGAGGGAATCTTAGAGG	1190	2211	TGATGAAAACAATAACAGGCCCAAGCCTGTGGTATGATGTGCAACACTTGTAGACTCAGAA	2270
Db	1190	AGATTAAACAGATATGGCCATTTTCAGCCACCATCAATTTGTAGAGGGAATCTTAGAGG	1249	2270	TGATGAAAACAATAACAGGCCCAAGCCTGTGGTATGATGTGCAACACTTGTAGACTCAGAA	2329
QY	1191	TTAAACATCATCCAGATGACAGCGTCTGATGCGGTTGCCATGGCCTGAAAGCTCCCTAA	1250	2271	AAAATACTACTCTCATAAATGGGTGGAGTATTTTGGTGACACCTTACTTTTGTCTGGCTG	2330
Db	1250	TTAAACATCATCCAGATGACAGCGTCTGATGCGGTTGCCATGGCCTGAAAGCTCCCTAA	1309	2330	AAAATACTACTCTCATAAATGGGTGGAGTATTTTGGTGACACCTTACTTTTGTCTGGCTG	2389
QY	1251	TAGACTTTGTGTCAGCTGCCAGGGAGCATTTCCACGGAGGTTCTGTACCATCATTTCTG	1310	2331	AGTGAAGGAATGATATTCATATATTCATTTATTTCCATGGACATTTAGTTAGTGTCTTTTA	2390
Db	1310	TAGACTTTGTGTCAGCTGCCAGGGAGCATTTCCACGGAGGTTCTGTACCATCATTTCTG	1369	2390	AGTGAAGGAATGATATTCATATATTCATTTATTTCCATGGACATTTAGTTAGTGTCTTTTA	2449
QY				2391	TATACCGGATGATGCTGAGTGACACTCTTGTGTATATTTCCAAAATTTTGTATAGTCTG	2450



Db 1070 CCAGACCTTCAAAACCCACCCCTCTCTTTAGCACTACTCTAAATCTTATGATCAAAACA 1129  
 Qy 1075 ----AGGACCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCC 1130  
 Db 1130 CCCAGGACCTTACTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCC 1189  
 Qy 1131 AGATTAAAGATATGGCAGCTTTCAAGCCACCATCACAAATCTAGAGGGAATCTTAGAGG 1190  
 Db 1190 AGATTAAAGATATGGCAGCTTTCAAGCCACCATCACAAATCTAGAGGGAATCTTAGAGG 1249  
 Qy 1191 TTAACATCATCCAGATGACAGACGCTCTGATGCCGGTGCATGGCTGAAAGCTCCCTAA 1250  
 Db 1250 TTAACATCATCCAGATGACAGACGCTCTGATGCCGGTGCATGGCTGAAAGCTCCCTAA 1309  
 Qy 1251 TAGACTTTGTCTGACCTGCGCAAGGAGCAITTCOCAGGAGTCTGTACCATCATTTCTG 1310  
 Db 1310 TAGACTTTGTCTGACCTGCGCAAGGAGCAITTCOCAGGAGTCTGTACCATCATTTCTG 1369  
 Qy 1311 ACCCCACCTGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGT 1370  
 Db 1370 ACCCCACCTGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGT 1429  
 Qy 1371 GTCTGCTGACTGTGAGACGAACTTCAATGGGTCTGGGACGTACTGTGTGAACTCACCC 1430  
 Db 1430 GTCTGCTGACTGTGAGACGAACTTCAATGGGTCTGGGACGTACTGTGTGAACTCACCC 1489  
 Qy 1431 TGGGGATGACCAAGCCTGGCTCTCAGAGCACCTGATTTCTGTTCTGACAGAGCC 1490  
 Db 1490 TGGGGATGACCAAGCCTGGCTCTCAGAGCACCTGATTTCTGTTCTGACAGAGCC 1549  
 Qy 1491 CAGCTCGCTTTAAGGATGGCAACAGTGCCTGATCTCCGTGGCTGCTGGCCATAT 1550  
 Db 1550 CAGCTCGCTTTAAGGATGGCAACAGTGCCTGATCTCCGTGGCTGCTGGCCATAT 1609  
 Qy 1551 TTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACACAGGAATACAAACCAATGAAA 1610  
 Db 1610 TTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACACAGGAATACAAACCAATGAAA 1669  
 Qy 1611 ATAGTCTGGGAATGTGGTGAGAACCAAGGCTGAGTGTCTTCTCAACCGTGCAAG 1670  
 Db 1670 ATAGTCTGGGAATGTGGTGAGAACCAAGGCTGAGTGTCTTCTCAACCGTGCAAG 1729  
 Qy 1671 CGGTGCTTCCCGGGAACCAAGGATCGCTACTCAAAAAACCAAGGAATTTAAAG 1730  
 Db 1730 CGGTGCTTCCCGGGAACCAAGGATCGCTACTCAAAAAACCAAGGAATTTAAAG 1789  
 Qy 1731 GAGTTTCTTAAATTTGACCTTGTCTGAACTGCTTCTGAGCTCACTTTTCAAGTGTGAGA 1790  
 Db 1790 GAGTTTCTTAAATTTGACCTTGTCTGAACTGCTTCTGAGCTCACTTTTCAAGTGTGAGA 1849  
 Qy 1791 TGTGCTGGAGTGGCTAATACCTTTTCTTAAAGATTTTGTAAATAGATATTGTGG 1850  
 Db 1850 TGTGCTGGAGTGGCTAATACCTTTTCTTAAAGATTTTGTAAATAGATATTGTGG 1909  
 Qy 1851 TTTGGGGAAGTGAATTTTATAGGTTAAATGTCATTTTAGAGATGGGAGGGATTA 1910  
 Db 1910 TTTGGGGAAGTGAATTTTATAGGTTAAATGTCATTTTAGAGATGGGAGGGATTA 1969  
 Qy 1911 TACTGAGGAGCTTACAGCCATGTGTGAACTGATAAAGCAACTTAGCAAGCTTCTT 1970  
 Db 1970 TACTGAGGAGCTTACAGCCATGTGTGAACTGATAAAGCAACTTAGCAAGCTTCTT 2029  
 Qy 1971 TTAATTAATTTTATGTTTCACTTATAAGTCTTAGTAACTAGTAGAGTAAGCAACTG 2030  
 Db 2030 TTAATTAATTTTATGTTTCACTTATAAGTCTTAGTAACTAGTAGAGTAAGCAACTG 2089  
 Qy 2031 TGTCCGAGAGTAAAGGAGAGAGTACTATTGATAGAGCTTAACCCAGGTTAACTGCAA 2090  
 Db 2090 TGTCCGAGAGTAAAGGAGAGAGTACTATTGATAGAGCTTAACCCAGGTTAACTGCAA 2149  
 Qy 2091 GAAGAGGCGGATCTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTT 2150  
 Db 2150 GAAGAGGCGGATCTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTT 2209

Qy 2151 GTTCTTAAGATCATGTTCCAAAGCTAACTGAATCCCACTTCAATACACACTCATGAACCTCC 2210  
 Db 2210 GTTCTTAAGATCATGTTCCAAAGCTAACTGAATCCCACTTCAATACACACTCATGAACCTCC 2269  
 Qy 2211 TGATGGAAACAATAACAGGCCCAAGCCTGTGTATGATGTGCACACTTCTAGACTCAGAA 2270  
 Db 2270 TGATGGAAACAATAACAGGCCCAAGCCTGTGTATGATGTGCACACTTCTAGACTCAGAA 2329  
 Qy 2271 AAAATACTACTCTCAATAATGGTGGAGATTTTGGTGAAACCTACTTTTCTTGGCTG 2330  
 Db 2330 AAAATACTACTCTCAATAATGGTGGAGATTTTGGTGAAACCTACTTTTCTTGGCTG 2389  
 Qy 2331 AGTGAAGGAATGATATTCATATATTCATTTATTCATGGACATTTAGTTAGTGTCTTTTA 2449  
 Db 2390 AGTGAAGGAATGATATTCATATATTCATTTATTCATGGACATTTAGTTAGTGTCTTTTA 2449  
 Qy 2391 TATACAGGATGATGCTGAGTGACACTTGTGTATTTCCAAATTTTCTATAGTCG 2450  
 Db 2450 TATACAGGATGATGCTGAGTGACACTTGTGTATTTCCAAATTTTCTATAGTCG 2509  
 Qy 2451 CTGCACATATTTGAAATCAAAATATTAAGACTTTCCAAAAATTTGGTCCCTGTTTCA 2510  
 Db 2510 CTGCACATATTTGAAATC-ATATATTAAGACTTTCCAAAGATGAGTCCCTGTTTCA 2568  
 Qy 2511 TGGCAACTTGATCAGTAAGGATTTCCCTCTCTTTGGAACTAAACCATTTACTATATGT 2570  
 Db 2569 TGGCAACTTGATCAGTAAGGATTTCCCTCTCTTTGTAACATAAACCATTTACTATATGT 2628  
 Qy 2571 TAGACAGACATTTTCTTCT 2624  
 Db 2629 TAGACATGACATTTCTTCT 2683

RESULT 10

ACD68613  
 ID ACD68613 standard; cDNA; 2683 BP.

XX ACD68613;

XX 17-SEP-2003 (first entry)

XX Novel human secreted and transmembrane protein PRO9925 cDNA.

XX Human; secreted and transmembrane protein; PRO; cytosolic;  
 XX antiarthritic; osteopathic; gene therapy; TNF-Agonist-Alpha;  
 XX chondrocyte stimulator; pericyte stimulator; fibroblast modulator;  
 XX pharmaceutical; diagnostic; biosensor; bioresactor; tumour; lung tumour;  
 XX colon tumour; breast tumour; prostate tumour; rectal tumour;  
 XX liver tumour; bone disorder; cartilage disorder; sports injury;  
 XX arthritis; wound; gene; ss.

XX Homo sapiens.

XX US2003045687-AI.

XX 06-MAR-2003.

XX 12-AUG-2002; 2002US-00218631.

XX 01-JUN-2001; 2001WO-US017800.

XX 29-JUN-2001; 2001WO-US021066.

XX 09-APR-2002; 2002US-00119480.

XX (SETH ) GENENTECH INC.

XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;  
 PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WJ;  
 XX WPI; 2003-512315/48.  
 XX P-PSDB; ABO33725.

XX New genes, and its encoded secreted and transmembrane polypeptides,

useful for stimulating Tumor Necrosis Factor alpha, or chondrocyte or pericyte proliferation, especially for treating lung tumors, arthritis or wounds in a mammal.

Claim 2; Fig 41; 314pp; English.

The invention describes an isolated nucleic acid molecule comprising a sequence with at least 80% identity to: (a) a nucleotide encoding any of 122 PRO (secreted and transmembrane) polypeptides whose sequences are fully defined in the specification; or (b) any of 122 nucleotide sequences having e.g. 4834, 2504 or 1759 bp fully defined in the specification; or the full length coding sequence of any of these 122 nucleotide sequences. The PRO polypeptides or polynucleotides are useful as pharmaceuticals, diagnostics, biosensors or bioreactors. These are particularly useful for detecting tumours (e.g. lung tumour, colon tumour, breast tumour, prostate tumour, rectal tumour, or liver tumour) in a mammal, for stimulating the release of TNF-alpha from human blood, for stimulating the proliferation or differentiation of chondrocyte cells, for stimulating proliferation of pericyte cells, or for modulating normal human dermal fibroblast proliferation. The PRO nucleic acid or polypeptide is also useful for treating tumours or various bone and/or cartilage disorders (e.g. sports injuries or arthritis), or wounds. The PRO polypeptides are useful in drug screening, particularly as targets for therapeutic intervention in these diseases, and in the diagnostic determination of the presence of these diseases. The PRO polypeptides are also useful as molecular weight markers, or for chromosome identification. The PRO genes are useful as hybridisation probes, or for screening libraries of human cDNA, genomic DNA or mRNA. The PRO genes may also be used in gene therapy, particularly for replacing a defective gene. This sequence encodes a novel human secreted and transmembrane PRO polypeptide

Sequence 2683 BP; 751 A; 604 C; 596 G; 732 T; 0 U; 0 Other;

Query Match 93.4%; Score 2485; DB 7; Length 2683; Best Local Similarity 97.8%; Pred. No. 0; Matches 2576; Conservative 0; Mismatches 20; Indels 39; Gaps 4;

PT	568	GATGAATTTATCTACGCTTCTCCACACACTTGGTCACTATTTCCAGAAATTTGGGACGAT	627
PT	590	GATGAATTTATCTACGCTTCTCCACACACTTGGTCACTATTTCCAGAAATTTGGGACGAT	649
PS			
XX			
XX			
CC	628	GTTGAGTGGAGTTTCTGTGGAACACAGCCAAATGTGACACTTGGGCTCAACTCATGGAAG	687
CC	650	GTTGAGTGGAGTTTCTGTGGAACACAGCCAAATGTGACACTTGGGCTCAACTCATGGAAG	709
CC			
CC	688	TGACTGTCTTACAGAGACATGACGGGATATGTTCCCATCCACACAGTGAAGATGTGT	747
CC	710	TGACTGTCTTACAGAGACATGACGGGATATGTTCCCATCCACACAGTGAAGATGTGT	769
CC			
CC	748	ACGTGGTAAACAGATCAGATTCCTGTGTTTGTGACTATGTTCCAGAGAACGATCGAAAT	807
CC	770	ACGTGGTAAACAGATCAGATTCCTGTGTTTGTGACTATGTTCCAGAGAACGATCGAAAT	829
CC			
CC	808	CATCCGAGAAACCTTCC-CAAAGATCTCCCATATGTTGATGCTCTGATTCATGATC	866
CC	830	CATCCGAGAAACCTTCC-CAAAGATCTCCCATATGTTGATGCTCTGATTCATGATC	889
CC			
CC	867	CTAGCCACTTCTCAATTAATCTCAATTAATCACTAAGTGGAGCTTCGGGATATATCTG	926
CC	890	CTAGCCACTTCTCAATTAATCTCAATTAATCACTAAGTGGAGCTTCGGGATATATCTG	949
CC			
CC	927	GCCTGTTGTTTCCACCAATCATCTGTAATCACACGTATGCTCAATGGAACCTTCA	986
CC	950	GCCTGTTGTTTCCACCAATCATCTGTAATCACACGTATGCTCAATGGAACCTTCA	1009
CC			
CC	987	GCCTTAACCTCACTGTGAAAGCTGAGCACAGGACCTTGTCCGCCACCCACCCAC	1046
CC	1010	GCCTTAACCTCACTGTGAAAGCTGAGCACAGGACCTTGTCCGCCACCCACCCAC	1069
CC			
CC	1047	CCAGACCTTCAAAACCCACCCCTCTTT-----	1074
CC	1070	CCAGACCTTCAAAACCCACCCCTCTTTTAGCAACTACTCTAAATCTTATGATCAAAACA	1129
CC			
CC	1075	---AGGACCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTCATGAACCTGCC	1130
CC	1130	CCCCAGGACCTACTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTCATGAACCTGCC	1189
CC			
CC	1131	AGATTACAGATATGGGCACTTTCAAGCCACCATCAATTTAGAGGGAATCTTAGAGG	1190
CC	1190	AGATTACAGATATGGGCACTTTCAAGCCACCATCAATTTAGAGGGAATCTTAGAGG	1249
CC			
CC	1191	TAAACATCATCCAGATGACAGACGCTGATGCGGTGCCATGCGCTGAAAGCTCCCTAA	1250
CC	1250	TAAACATCATCCAGATGACAGACGCTGATGCGGTGCCATGCGCTGAAAGCTCCCTAA	1309
CC			
CC	1251	TAGACTTTGTGCTGACCTGCCAGGGAGCATTCACCGGAGGTCTGTACCATCATTTCTG	1310
CC	1310	TAGACTTTGTGCTGACCTGCCAGGGAGCATTCACCGGAGGTCTGTACCATCATTTCTG	1369
CC			
CC	1311	ACCCACCTGGGAGATCACCCAGAACACAGCTGTGACCCCTGTGGATGTGGATGATGT	1370
CC	1370	ACCCACCTGGGAGATCACCCAGAACACAGCTGTGACCCCTGTGGATGTGGATGATGT	1429
CC			
CC	1371	GTCTGTGACTGTGAGACGAACTTCAATGGGTCTGGAGCTGTGTGAACTCATCC	1430
CC	1430	GTCTGTGACTGTGAGACGAACTTCAATGGGTCTGGAGCTGTGTGAACTCATCC	1489
CC			
CC	1431	TGGGGATGACACAAGCCTTGGCTCTCACAGACACCTGATTTCTGTTCTTCGACAGAC	1490
CC	1490	TGGGGATGACACAAGCCTTGGCTCTCACAGACACCTGATTTCTGTTCTTCGACAGAC	1549
CC			
CC	1491	CAGCCTCGCCTTTAAGATGGGAAAAAGTGCCTGTATCTCCGTGGCTGTGGCCATAT	1550
CC	1550	CAGCCTCGCCTTTAAGATGGGAAAAAGTGCCTGTATCTCCGTGGCTGTGGCCATAT	1609
CC			
CC	1551	TTCTCACTGTGATCTCCCTCTTGGTGTACAAAAACACAGGAATACACCCCAATGAAA	1610
CC	1610	TTCTCACTGTGATCTCCCTCTTGGTGTACAAAAACACAGGAATACACCCCAATGAAA	1669

QY	28	AACCTTGGTGGCTGGTCCGTGAGAAATTCAGCATGGAATGTCTTACTATTTCTGGGAT	87
DB	50	ACCTTGGTGGCTGGTCCGTGAGAAATTCAGCATGGAATGTCTTACTATTTCTGGGAT	109
QY	88	TTCTGCTCTGGCTGCAAGATGTCACCTTGATGCGCCAAACGATTTTCATGATGTGTGG	147
DB	110	TTCTGCTCTGGCTGCAAGATGTCACCTTGATGCGCCAAACGATTTTCATGATGTGTGG	169
QY	148	GCAATGAAGACCTTCTGCTTACATGAGGAGCAATCAATTAATGGCTGGTCTTCTG	207
DB	170	GCAATGAAGACCTTCTGCTTACATGAGGAGCAATCAATTAATGGCTGGTCTTCTG	229
QY	208	ATGAAATGACTGGAATGAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGTGA	267
DB	230	ATGAAATGACTGGAATGAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGTGA	289
QY	268	AAAACTCCTGGAAGGAGGCGGTGTGACGGCGGTCTGACCCAGTCACTCACCCCTCG	327
DB	290	AAAACTCCTGGAAGGAGGCGGTGTGACGGCGGTCTGACCCAGTCACTCACCCCTCG	349
QY	328	TGGGCTCAATATTAACATTTGCGGTGAACCTGATATTTCCCTAGATGCCAAAGGAAGT	387
DB	350	TGGGCTCAATATTAACATTTGCGGTGAACCTGATATTTCCCTAGATGCCAAAGGAAGT	409
QY	388	CCATGCAACATAGTCTATGAGAGAACCTGCAAAATGAGGCTGGTTATCTGCTGATC	447
DB	410	CCATGCAACATAGTCTATGAGAGAACCTGCAAAATGAGGCTGGTTATCTGCTGATC	469
QY	448	CATATGTTTAACTGACAGCATGTGTGAGGAGCATGACGGGAAAAATGGCACCGCC	507
DB	470	CGTATGTTTAACTGACAGCATGTGTGAGGAGCATGACGGGAAAAATGGCACCGCC	529
QY	508	AAAGCCATCAATAGCTCTTCCCTGATGGGAAACCTTTTCTCACCAACCCCGGATGAGAA	567
DB	530	AAAGCCATCAATAGCTCTTCCCTGATGGGAAACCTTTTCTCACCAACCCCGGATGAGAA	589







QY 208 ATGAAATGACTGGATGAAATCTTACCAGTGTGGAAGCGGGAGACATGAGGTGA 267  
Db 230 ATGAAATGACTGGATGAAATCTTACCAGTGTGGAAGCGGGAGACATGAGGTGA 289  
QY 268 AAAATCTCTGGAAGGAGCCGTGTGACGCGGTCTGACCACTGAGTCACTCACCAGCCCTCG 327  
Db 290 AAAATCTCTGGAAGGAGCCGTGTGACGCGGTCTGACCACTGAGTCACTCACCAGCCCTCG 349  
QY 328 TGGGCTCAAAATATACATTTGCGGTGAACCTGATATTCCTAGATGCCCCAAAAGGAGATG 387  
Db 350 TGGGCTCAAAATATACATTTGCGGTGAACCTGATATTCCTAGATGCCCCAAAAGGAGATG 409  
QY 388 CCAATGGCAACATAGTCTATGAGAGAACTGCAGAAATGAGCTGGTTTATCTGCTGATC 447  
Db 410 CCAATGGCAACATAGTCTATGAGAGAACTGCAGAAATGAGCTGGTTTATCTGCTGATC 469  
QY 448 CATATGTTTACAACCTGGACAGCATGGTCAAGAGACAGTGACGGGAAATGGCACCGGCC 507  
Db 470 CGTATGTTTACAACCTGGACAGCATGGTCAAGAGACAGTGACGGGAAATGGCACCGGCC 529  
QY 508 AAAGCCATCAAACTCTTCCCTGATGGGAAACCTTTTCTCACCACCCCGGATGGAGAA 567  
Db 530 AAAGCCATCAAACTCTTCCCTGATGGGAAACCTTTTCTCACCACCCCGGATGGAGAA 589  
QY 568 GATGAAATTTCAATCTACGTCTTCCACACATTTGTCAGTATTTCCAGAAATGGACAGAT 627  
Db 590 GATGAAATTTCAATCTACGTCTTCCACACATTTGTCAGTATTTCCAGAAATGGACAGAT 649  
QY 628 GTTCAGTGAGAGTTTCTGTGAACACAGCAATGTGACACTTGGGCTCAACTCATGGAAG 687  
Db 650 GTTCAGTGAGAGTTTCTGTGAACACAGCAATGTGACACTTGGGCTCAACTCATGGAAG 709  
QY 688 TGACTGTCTACAGAGACATGGAACGGGATATGTTCCCATGCGCAACAGTGAAGATGTGT 747  
Db 710 TGACTGTCTACAGAGACATGGAACGGGATATGTTCCCATGCGCAACAGTGAAGATGTGT 769  
QY 748 ACGTGTGAACAGATCAGATTCCTGTGTTGTGACTATGTTCCATGCGCAACAGTGAAGAT 807  
Db 770 ACGTGTGAACAGATCAGATTCCTGTGTTGTGACTATGTTCCATGCGCAACAGTGAAGAT 829  
QY 808 CATCCGAGAAACCTTCC- CAAAGATCTCCCATATGTTGATGTCCTGATTCATGATC 866  
Db 830 CATCCGAGAAACCTTCC- CAAAGATCTCCCATATGTTGATGTCCTGATTCATGATC 889  
QY 867 CTAGCCATCTCTCAATATTTCTACCAATTAACAGTGAAGTTCGGGATTAATCTG 926  
Db 890 CTAGCCATCTCTCAATATTTCTACCAATTAACAGTGAAGTTCGGGATTAATCTG 949  
QY 927 GCCTGTTGTTTCCACCAATCATACTGTGAATCAACGATATGTCATGGAACCTTCA 986  
Db 950 GCCTGTTGTTTCCACCAATCATACTGTGAATCAACGATATGTCATGGAACCTTCA 1009  
QY 987 GCCTTAACCTCACTGTGAAAGCTGCAGACACAGACCTTGTCCGCCACCGCCACACAC 1046  
Db 1010 GCCTTAACCTCACTGTGAAAGCTGCAGACACAGACCTTGTCCGCCACCGCCACACAC 1069  
QY 1047 CCAGACCTTCAAAACCCACCCCTTCTTTAGCAACTACTCTAAATCTTATGATTCAAACA 1074  
Db 1070 CCAGACCTTCAAAACCCACCCCTTCTTTAGCAACTACTCTAAATCTTATGATTCAAACA 1129  
QY 1075 ----AGGACCTGTGGTGAACACCCCTGGAGCTGAGTAGATTCCTGATGAAACCTGCC 1130  
Db 1130 CCCAGAGACCTTACTGGGTGACAAACCCCTGGAGCTGAGTAGATTCCTGATGAAACCTGCC 1189  
QY 1131 AGATTAACAGATATGGCCACTTTCAAGCCACCATCAAAATGTAGAGGAATCTTAGAGG 1190  
Db 1190 AGATTAACAGATATGGCCACTTTCAAGCCACCATCAAAATGTAGAGGAATCTTAGAGG 1249  
QY 1191 TTAACATCATCCAGATGACAGACGTCTGTATGCGGTTGCCATGGCTGAAAGCTCCCTAA 1250  
Db 1250 TTAACATCATCCAGATGACAGACGTCTGTATGCGGTTGCCATGGCTGAAAGCTCCCTAA 1309

QY 1251 TAGACTTTGTCTGTGACCTGCCAAGGAGCAATTCACCGAGGTCTTACCATCAATTTCTG 1310  
Db 1310 TAGACTTTGTCTGTGACCTGCCAAGGAGCAATTCACCGAGGTCTTACCATCAATTTCTG 1369  
QY 1311 ACCCCACCTGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGATGT 1370  
Db 1370 ACCCCACCTGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGATGT 1429  
QY 1371 GTCTGTGACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCC 1430  
Db 1430 GTCTGTGACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCC 1489  
QY 1431 TGGGGATGACACAAAGCTGGCTCTCACGAGCACCTGATTTCTTCTCTGACAGAGACC 1490  
Db 1490 TGGGGATGACACAAAGCTGGCTCTCACGAGCACCTGATTTCTTCTCTGACAGAGACC 1549  
QY 1491 CAGCCCTCGCCCTTAAAGGATGGCAACAGTGCCTGATCTCCGTTGGCTGTGGCCATAT 1550  
Db 1550 CAGCCCTCGCCCTTAAAGGATGGCAACAGTGCCTGATCTCCGTTGGCTGTGGCCATAT 1609  
QY 1551 TTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACACAAAGAAATACAAACCAATAGAAA 1610  
Db 1610 TTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACACAAAGAAATACAAACCAATAGAAA 1669  
QY 1611 ATAGTCTGGGAATGTGTGAGAACAAAGGCTGAGTGTCTTCTCAACCGTGCAAAAG 1670  
Db 1670 ATAGTCTGGGAATGTGTGAGAACAAAGGCTGAGTGTCTTCTCAACCGTGCAAAAG 1729  
QY 1671 CCGTGTCTTCCCGGAAACCCAGGAAAGGATCGCTACTCAAAAACCAAGAAATTAAG 1730  
Db 1730 CCGTGTCTTCCCGGAAACCCAGGAAAGGATCGCTACTCAAAAACCAAGAAATTAAG 1789  
QY 1731 GAGTTTCTTAAATTTCCGACCTTGTTCGAAAGCTACATTTTCAGTGCCATGTAGTGAGA 1790  
Db 1790 GAGTTTCTTAAATTTCCGACCTTGTTCGAAAGCTACATTTTCAGTGCCATGTAGTGAGA 1849  
QY 1791 TGTGCTGAGTGGCTATTAACCTTTTTCCTAAAGATTTTGTAAATAGATATTTGG 1850  
Db 1850 TGTGCTGAGTGGCTATTAACCTTTTTCCTAAAGATTTTGTAAATAGATATTTGG 1909  
QY 1851 TTTCCGGAGCTGAATTTTTTATAGTTAAATGTCTATTTAGATGGGAGAGGATTA 1910  
Db 1910 TTTCCGGAGCTGAATTTTTTATAGTTAAATGTCTATTTAGATGGGAGAGGATTA 1969  
QY 1911 TACTGCGAGGAGCTTCAGCCATGTTGTGAACCTCATAAAGCAACTTAGCAAGCTCTT 1970  
Db 1970 TACTGCGAGGAGCTTCAGCCATGTTGTGAACCTCATAAAGCAACTTAGCAAGCTCTT 2029  
QY 1971 TTTCAATTTTATGTTTCACTTATAAGTCTTAGGTAACTAGTAGATAGAAACACTG 2030  
Db 2030 TTTCAATTTTATGTTTCACTTATAAGTCTTAGGTAACTAGTAGATAGAAACACTG 2089  
QY 2031 TGTCCCGAGAGTAGGAGAGAGCTACTATTTAGTAGGCTTAACCCAGGTTAACTGCAA 2090  
Db 2090 TGTCCCGAGAGTAGGAGAGAGCTACTATTTAGTAGGCTTAACCCAGGTTAACTGCAA 2149  
QY 2091 GAAGAGCGGATACTTTTCAGCTTTTCCATGCTATGATGATGATGATGATGATGATGAT 2150  
Db 2150 GAAGAGCGGATACTTTTCAGCTTTTCCATGCTATGATGATGATGATGATGATGATGAT 2209  
QY 2151 GTTTCTAAGATCAATGTTTCCAGCTAACTGAATCCACTTCAATACACACTCATGAATCC 2210  
Db 2210 GTTTCTAAGATCAATGTTTCCAGCTAACTGAATCCACTTCAATACACACTCATGAATCC 2269  
QY 2211 TGATGGAACAAATACAGGCCCAAGCCCTGTGGTATGATGATGATGATGATGATGATGAT 2270  
Db 2270 TGATGGAACAAATACAGGCCCAAGCCCTGTGGTATGATGATGATGATGATGATGATGAT 2329  
QY 2271 AAAATACTACTCTCATAAATGGGTGGGATTTTGGTGACAACTACTTTGCTGGCTG 2330  
Db 2330 AAAATACTACTCTCATAAATGGGTGGGATTTTGGTGACAACTACTTTGCTGGCTG 2389  
QY 2331 AGTGAAGGAATGATATTCATATATATTCATTTTCCATGGCAATTTAGTTAGTCTTTTA 2390









QY	508	AAGCCCATATAACGCTTCCCTGTATGGGAAACCTTTTCCTCACACCCTCGGATGGABRA	567
Db	530	AAAAGCATTATAACGCTTCCCTGTATGGGAAACCTTTTCCTCACACCCTCGGATGGAGAA	589
QY	568	GATGGAAATTCATCTACGTCTTCCACACACTTGGTGAGTAATTTCCAGAAATTTGGGACGAT	627
Db	590	GATGGAAATTCATCTACGTCTTCCACACACTTGGTGAGTAATTTCCAGAAATTTGGGACGAT	649
QY	628	GTTCAGTGAGAGTTTCTGTGAACAACAGCCAAATGTGACACTTTGGGCCCTCAACTCATGGAAG	687
Db	650	GTTCAGTGAGAGTTTCTGTGAACAACAGCCAAATGTGACACTTTGGGCCCTCAACTCATGGAAG	709
QY	688	TGACTGTCTACAGAGACATCGACGGGCATATGTTCCCATCGCACAAAGTGAAGAATGTGT	747
Db	710	TGACTGTCTACAGAGACATCGACGGGCATATGTTCCCATCGCACAAAGTGAAGAATGTGT	769
QY	748	ACGTGGTAAACAGATCAGATTCCCTGTGTTGTGACTATGTTCCAGAGAAACGATCGAAATTT	807
Db	770	ACGTGGTAAACAGATCAGATTCCCTGTGTTGTGACTATGTTCCAGAGAAACGATCGAAATTT	829
QY	808	CATCGGAGAAACCTTCC-CHAAAGATCCCCCATATGTTGTGATGCTCTGATTCATGATC	866
Db	830	CATCGGAGAAACCTTCCCAAAGATCTCCCACATATGTTGTGATGCTCTGATTCATGATC	889
QY	867	CTAGCCACTTCTCAATTAATTTCTACCATTAACACTAAGTGAGCTTCGGGGATAATACTG	926
Db	890	CTAGCCACTTCTCAATTAATTTCTACCATTAACACTAAGTGAGCTTCGGGGATAATACTG	949
QY	927	GCCTGTTTTTCCACCAATCATACTGTAATTCACACATGATGCTGTAAGAACCTTCA	986
Db	950	GCCTGTTTTTCCACCAATCATACTGTAATTCACACATGATGCTGTAAGAACCTTCA	1009
QY	987	GCCTTAAACCTCACTGTGAAGCTGCAGACACAGGACCTTGTCCGCCACCGCACCCAC	1046
Db	1010	GCCTTAAACCTCACTGTGAAGCTGCAGACACAGGACCTTGTCCGCCACCGCACCCAC	1069
QY	1047	CCAGACCTTCAAAACCCACCCCTTCTTT-----	1074
Db	1070	CCAGACCTTCAAAACCCACCCCTTCTTTTAGCAACTACTCTAAAAATCTTATGATTCAAACA	1129
QY	1075	----AGGACCTGCTGTGACAAACCCCTGGAGCTGAGTAGGATTCCTCATGAACACTGCC	1130
Db	1130	CGCCAGGACCTACTGTGTGACAAACCCCTGGAGCTGAGTAGGATTCCTCATGAACAACTGCC	1189
QY	1131	AGATTAAACAGATATGCCACTTTTCAAGCCACCATCACAAATTTGTAGAGGGAATCTTTAGAGG	1190
Db	1190	AGATTAAACAGATATGSCACTTTTCAAGCCACCATCACAAATTTGTAGAGGGAATCTTTAGAGG	1249
QY	1191	TTAAACATCATCCAGATGACAGAGCTGATGCGCGGTGCGATGGCTGAAAGCTCCCTAA	1250
Db	1250	TTAAACATCATCCAGATGACAGAGCTGATGCGCGGTGCGATGGCTGAAAGCTCCCTAA	1309
QY	1251	TAGACTTCTGCTGACCTGCCAGGAGGACATTTCCACGGAGGTCTGTACCAATCTTTCTG	1310
Db	1310	TAGACTTCTGCTGACCTGCCAGGAGGACATTTCCACGGAGGTCTGTACCAATCTTTCTG	1369
QY	1311	ACCCCACTCGGAGATCACCCGAAACACAGTCTGCAGCCCTGTGAGATGFGGATGAGATGT	1370
Db	1370	ACCCCACTCGGAGATCACCCGAAACACAGTCTGCAGCCCTGTGAGATGFGGATGAGATGT	1429
QY	1371	GTCTGTGACTGTGAGACGAACTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCC	1430
Db	1430	GTCTGTGACTGTGAGACGAACTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCC	1489
QY	1431	TGGGGGATGACACAAGCCTGGCTCTCACAGACACCCTGATTTCTGTTCTCTGCAGAGACC	1490
Db	1490	TGGGGGATGACACAAGCCTGGCTCTCACAGACACCCTGATTTCTGTTCTCTGCAGAGACC	1549
QY	1491	CAGCTCGCCCTTTAAGATGGGAAAACAGTCCCTGATCTCCGTTGGCTCTGGCCATAT	1550
Db	1550	CAGCTCGCCCTTTAAGATGGGAAAACAGTCCCTGATCTCCGTTGGCTCTGGCCATAT	1609

## RESULT 15

ABT43902  
XX ID ABT43902 standard; cDNA; 2683 BP.  
XX AC ABT43902;  
XX DT 16-OCT-2003 (first entry)  
XX DE Human membrane bound receptor/protein PRO9925 cDNA sequence.  
XX KW Human; PRO; membrane bound protein; membrane bound receptor;  
KW cell proliferation; cell migration; cell differentiation;  
KW mitogenic factor; survival factor; cytotoxic factor;  
KW differentiation factor; neurotrophic factor; hormone; cell receptor;  
KW receptor-ligand interaction; cytoskeletal; chondrocyte; tumour; gene; ss.  
XX OS Homo sapiens.

XX US 2003065147-A1.  
XX PN 03-APR-2003.  
XX PD 29-AUG-2002; 2002US-00232224.  
XX PF 28-JUL-1999; 99US-0146222P.  
XX PR 24-FEB-2000; 2000WO-US005004.  
XX PR 02-MAR-2000; 2000WO-US005841.  
XX PR 01-JUN-2001; 2001WO-US017800.  
XX PR 29-JUN-2001; 2001WO-US021066.  
XX PR 09-APR-2002; 2002US-00119480.  
XX PA (GETH) GENENTECH INC.

XX PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;  
PI Gramaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;  
XX P-PSDB; ABU72078.  
XX WPI; 2003-522018/49.

XX PT One hundred and twenty two nucleic acids encoding PRO polypeptides,  
PT useful for the manufacture of a medicament for diagnosing or treating  
PT tumor.

XX PS Claim 2; Fig 41; 315pp; English.

XX This invention relates to one hundred and twenty two novel nucleic acids  
XX encoding human PRO membrane bound proteins or receptors. Extracellular  
XX proteins play important roles in the formation, differentiation and  
XX maintenance of multicellular organisms. The fate of many individual cells  
XX (for example proliferation, migration or differentiation) is typically  
XX governed by information received from other cells and the immediate  
XX environment. The information is often transmitted by secreted  
XX polypeptides (for example mitogenic factors, survival factors, cytotoxic  
XX factors, differentiation factors, neurotrophic factors and hormones) which are  
XX received and interpreted by diverse cell receptors or membrane bound  
XX proteins. These membrane bound proteins and receptors may be of use as  
XX pharmaceutical and diagnostic agents, such as in the blocking of receptor  
XX -ligand interactions. The current invention provides the amino acid  
XX sequences of novel human membrane bound receptors and proteins, along  
XX with the cDNA sequences encoding them. The novel proteins of the  
XX invention may have cytostatic activities through the stimulation of  
XX chondrocytes. The nucleic acids of the invention may be useful for the  
XX manufacture of a medicament for diagnosing or treating a tumour in a  
XX mammal. In addition, they may be useful for measuring or detecting the  
XX expression of a tumour associated gene. The present sequence is the cDNA  
XX sequence encoding a human PRO protein of the invention

XX SQ Sequence 2683 BP; 751 A; 604 C; 596 G; 732 T; 0 U; 0 Other;  
XX Query Match 93.4%; Score 2485; DB 8; Length 2683;  
XX Best Local Similarity 97.8%; Pred. No. 0;  
XX Matches 2576; Conservative 0; Mismatches 20; Indels 39; Gaps 4;

QY 28 AACCTTGGTCCCTGGCTCCGTGAGAAATTCAGCATGGAATGCTCTACTATTTCTCTGGAT 87  
DB 50 ACCTTGAGTGCCTGGCTCCGTGAGAAATTCAGCATGGAATGCTCTACTATTTCTCTGGAT 109  
QY 88 TTCTGCTCCTGGCTGCAAGATTCACCTGATGCGCCCAACGATTCATGATGCTGCTG 147  
DB 110 TTCTGCTCCTGGCTGCAAGATTCACCTGATGCGCCCAACGATTCATGATGCTGCTG 169  
QY 148 GCAATGAAAGACCTTCTGCTTACATGAGGAGCACAATCAATTAATGCTGCTCTG 207  
DB 170 GCAATGAAAGACCTTCTGCTTACATGAGGAGCACAATCAATTAATGCTGCTCTG 229  
QY 208 ATGAAATGACTGGAATGAAATCTACCCAGTGTGGAAGCGGGAGACATGAGTGA 267  
DB 230 ATGAAATGACTGGAATGAAATCTACCCAGTGTGGAAGCGGGAGACATGAGTGA 289  
QY 268 AAACTCCTGGAAGGAGCGCTGTGAGGCGGTCTGACAGTGAATCCAGAGTGAATG 327  
DB 290 AAACTCCTGGAAGGAGCGCTGTGAGGCGGTCTGACAGTGAATCCAGAGTGAATG 349  
QY 328 TGGGCTCAAAATATAACATTTGGGTGAACCTGATATCCCTAGATGCCAAAGGAGATG 387  
DB 350 TGGGCTCAAAATATAACATTTGGGTGAACCTGATATCCCTAGATGCCAAAGGAGATG 409  
QY 388 CCAATGGCAACATAGTCTATGAGAAAGAACTGCAGAAATGAGGCTGCTTATCTGCTGATC 447  
DB 410 CCAATGGCAACATAGTCTATGAGAAAGAACTGCAGAAATGAGGCTGCTTATCTGCTGATC 469  
QY 448 CATATGTTTACAACCTGGAAGGAGCGCTGTGAGGCGGTCTGACAGTGAATGCAACGCGCC 507  
DB 470 CGTATGTTTACAACCTGGAAGGAGCGCTGTGAGGCGGTCTGACAGTGAATGCAACGCGCC 529  
QY 508 AAAGCCATCAATAACGCTCTTCCCTGATGGAAACCTTTTCCCTACACACCCCGATGAGAA 567  
DB 530 AAAGCCATCAATAACGCTCTTCCCTGATGGAAACCTTTTCCCTACACACCCCGATGAGAA 589  
QY 568 GATGGAATTTTCACTACGCTCTTCCACACACCTTGCTGATTTTCCAGAAATTTGGACGAT 627  
DB 590 GATGGAATTTTCACTACGCTCTTCCACACACCTTGCTGATTTTCCAGAAATTTGGACGAT 649  
QY 628 GTTCAGTGAGAGTTTCTGTGAACAGCCATGACACCTTGGGCTCAACCTATGGAAG 687  
DB 650 GTTCAGTGAGAGTTTCTGTGAACAGCCATGACACCTTGGGCTCAACCTATGGAAG 709  
QY 688 TGACTGTCTACAGAAAGACATGGACGGCATATGTTCCCATCGCAAGTGAAGATGAT 747  
DB 710 TGACTGTCTACAGAAAGACATGGACGGCATATGTTCCCATCGCAAGTGAAGATGAT 769  
QY 748 ACGTGGTAAACAGATCAGATTCCTGTTGTGACTATGTTCCAGAAACGATCGAAAT 807  
DB 770 ACGTGGTAAACAGATCAGATTCCTGTTGTGACTATGTTCCAGAAACGATCGAAAT 829  
QY 808 CATCCGAGAAACCTTCC-CAAGATCTCCCATATATGTTTGTGATGCTGATGATGATC 866  
DB 830 CATCCGAGAAACCTTCC-CAAGATCTCCCATATATGTTTGTGATGCTGATGATGATC 889  
QY 867 CTAGCCACTTCCCTCAATTTATTCTACCAATTAACCTACAGTGAAGTGGGATTAATCTG 926  
DB 890 CTAGCCACTTCCCTCAATTTATTCTACCAATTAACCTACAGTGAAGTGGGATTAATCTG 949  
QY 927 GCTGTTGTTTCCCAACATCATATGTTGAATCAGATGATGCTCAATGGAACCTTCA 986  
DB 950 GCTGTTGTTTCCCAACATCATATGTTGAATCAGATGATGCTCAATGGAACCTTCA 1009  
QY 987 GCTTAACTCCTGTAAGAGTGCAGACAGGCTTGTCCGCAACGCCACCCAC 1046  
DB 1010 GCTTAACTCCTGTAAGAGTGCAGACAGGCTTGTCCGCAACGCCACCCAC 1069  
QY 1047 CCAGACCTTCAAAACCCACCCCTTCTTT----- 1074  
DB 1070 CCAGACCTTCAAAACCCACCCCTTCTTTTAGCAACTACTTAAATCTTATGATTTCAACA 1129  
QY 1075 ----AGGACCTGCTGTGACACACCCCTGAGTGAAGTATTCCTGATGAAATGCTGCC 1130



Db	1130	CCCCAGGACCTACTGGTGACAACCCCCCTGGAGCTGAGTAGGATCTCTGATGAAAACTGCC	1189
Qy	1131	AGATTAAACAGATATGGCCACTTTCAAGCCACCATCACAAATTCTAGAGGGAATCTTAGAGG	1190
Db	1190	AGATTAAACAGATATGGCCACTTTCAAGCCACCATCACAAATTGTAGAGGGAATCTTAGAGG	1249
Qy	1191	TTAAACATATCCAGATGACAGACGTCCTGATGCCGTGCCATCGGCTGAAAGCTCCCTAA	1250
Db	1250	TTAAACATATCCAGATGACAGACGTCCTGATGCCGTGCCATCGGCTGAAAGCTCCCTAA	1309
Qy	1251	TAGACTTTTGCTGTAAGCTGCTCAAGGAGCAATCCCAAGGAGTCTGTACAATCATTTCTG	1310
Db	1310	TAGACTTTTGCTGTAAGCTGCTCAAGGAGCAATCCCAAGGAGTCTGTACAATCATTTCTG	1369
Qy	1311	ACCCCACTCGGAGATCACCCAGAACACAGTCTGACAGCCCTGTGGATGGATGAGATGT	1370
Db	1370	ACCCCACTCGGAGATCACCCAGAACACAGTCTGACAGCCCTGTGGATGGATGAGATGT	1429
Qy	1371	GTCTGCTGACTGTGAGAGCAACTTCAATGGGTCTGGAGCTACTGTGTAACCTCACCC	1430
Db	1430	GTCTGCTGACTGTGAGAGCAACTTCAATGGGTCTGGAGCTACTGTGTAACCTCACCC	1489
Qy	1431	TGGGGGATGACAAGCCTGGCTCTACGAGACCCCTGATTTCTGTTCTCTGACAGAGACC	1490
Db	1490	TGGGGGATGACAAGCCTGGCTCTACGAGACCCCTGATTTCTGTTCTCTGACAGAGACC	1549
Qy	1491	CAGCCTCGCCTTTAAGGATGGCAACAGTGGCCTGATCTCCGTGGCTGCTTGGCCATAT	1550
Db	1550	CAGCCTCGCCTTTAAGGATGGCAACAGTGGCCTGATCTCCGTGGCTGCTTGGCCATAT	1609
Qy	1551	TTGTCACTGTGATCTCCCTCTTGGTGTAACAAAAACAAAGGAATACAAACCAATAGAAA	1610
Db	1610	TTGTCACTGTGATCTCCCTCTTGGTGTAACAAAAACAAAGGAATACAAACCAATAGAAA	1669
Qy	1611	ATAGTCTCTGGGAATGTGGTCAGAGCAAAAGCCTGAGTGTCTTCTCAACCGTGGAAAAG	1670
Db	1670	ATAGTCTCTGGGAATGTGGTCAGAGCAAAAGCCTGAGTGTCTTCTCAACCGTGGAAAAG	1729
Qy	1671	CCGTGTTCTCCCGGAAACAGCAAAAGGATCCGCTACTCAAAAACCAAGAAATTTAAAG	1730
Db	1730	CCGTGTTCTCCCGGAAACAGCAAAAGGATCCGCTACTCAAAAACCAAGAAATTTAAAG	1789
Qy	1731	GAGTTTCTTAAATTTGACCTTGTCTGAAGCTCACTTTTCAGTGCCAATGATGTGAGA	1790
Db	1790	GAGTTTCTTAAATTTGACCTTGTCTGAAGCTCACTTTTCAGTGCCAATGATGTGAGA	1849
Qy	1791	TGTCCTGGAGTGGCTATTAACCTTTTTTCTTAAAGATATATGTTAAATAGATTTGTGG	1850
Db	1850	TGTCCTGGAGTGGCTATTAACCTTTTTTCTTAAAGATATATGTTAAATAGATTTGTGG	1909
Qy	1851	TTTGGGGAAGTTGAATTTTTTATAGGTTAAATGTCAATTTAGAGATGGGAGAGGATTA	1910
Db	1910	TTTGGGGAAGTTGAATTTTTTATAGGTTAAATGTCAATTTAGAGATGGGAGAGGATTA	1969
Qy	1911	TACTCAGGCAGCTTCAGCCATGTGTGAAACTGATAAAGCACTTAGCAAGCTTCTT	1970
Db	1970	TACTCAGGCAGCTTCAGCCATGTGTGAAACTGATAAAGCACTTAGCAAGCTTCTT	2029
Qy	1971	TTCAATTTTTTTATGTTTCACTTATAAAGTCTTAGGTAACCTAGTAGATAGAAACACTG	2030
Db	2030	TTCAATTTTTTTATGTTTCACTTATAAAGTCTTAGGTAACCTAGTAGATAGAAACACTG	2089
Qy	2031	TGTCCCGAGATAGGAGAGAGCTACTATTGATTAGAGCTTAACCCAGGTTAACTGCAA	2090
Db	2090	TGTCCCGAGATAGGAGAGAGCTACTATTGATTAGAGCTTAACCCAGGTTAACTGCAA	2149
Qy	2091	GAAGAGCGGGATACCTTTTCAGCTTTCCTCACTGTAACCTGATGATAAAGCAATGATGCCA	2150
Db	2150	GAAGAGCGGGATACCTTTTCAGCTTTCCTCACTGTAACCTGATGATAAAGCAATGATGCCA	2209
Qy	2151	GTTCCTAAGATCATGTTCCAGCTAACCTGAATCCCACTTCAATACACATCATGAATCC	2210



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: September 10, 2004, 15:48:34 ; Search time 569.5 Seconds  
(without alignments)  
2996.857 Million cell updates/sec

Title: US-10-039-272-1  
Perfect score: 4850  
Sequence: 1 cggcagcgggcccgagga.....aaaaaaaaaaaaaaaaaaaaa 2661

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 2670352

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ n2p.model -DEV=xlp  
-Q=/cgn2\_1/USPTC\_spool\_p/US10039272/runat\_07092004\_144225\_21764/app\_query.fasta\_1.2823  
-DB=Published Applications AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blowsum62  
-TRANS=human40.cdi -LIST=45 -DOCALL=200 -THR SCORE=pct-THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10039272@cgn\_1\_1\_262@runat\_07092004\_144225\_21764  
-NCFU=6 -ICFU=3 -NO\_WMAP -LARGSEQ=0 -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXI=7

Database :

1:	/cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
13:	/cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/1/pubpaa/US10F_PUBCOMB.pep.*
17:	/cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
-----					

1	2999	61.8	560	9	US-09-735-705-225	Sequence 225, App
2	2999	61.8	560	9	US-09-850-716A-225	Sequence 225, App
3	2999	61.8	560	9	US-09-897-778-225	Sequence 225, App
4	2999	61.8	560	9	US-09-943-075A-6	Sequence 6, Appli
5	2999	61.8	560	12	US-10-007-700-225	Sequence 225, App
6	2999	61.8	560	12	US-10-463-106-2	Sequence 2, Appli
7	2999	61.8	560	13	US-10-039-272-2	Sequence 2, Appli
8	2999	61.8	560	14	US-10-117-982-225	Sequence 225, App
9	2999	61.8	560	15	US-10-295-027-1258	Sequence 1258, Ap
10	2999	61.8	560	15	US-10-313-986-225	Sequence 225, App
11	2999	61.8	560	16	US-10-309-290-152	Sequence 152, App
12	2999	61.8	560	16	US-10-408-765A-466	Sequence 466, App
13	2979	61.4	572	12	US-10-219-535-42	Sequence 42, Appl
14	2979	61.4	572	12	US-10-232-230-42	Sequence 42, Appl
15	2979	61.4	572	12	US-10-232-224-42	Sequence 42, Appl
16	2979	61.4	572	14	US-10-227-884-42	Sequence 42, Appl
17	2979	61.4	572	14	US-10-230-163-42	Sequence 42, Appl
18	2979	61.4	572	14	US-10-230-338-42	Sequence 42, Appl
19	2979	61.4	572	14	US-10-218-631-42	Sequence 42, Appl
20	2979	61.4	572	14	US-10-230-414-42	Sequence 42, Appl
21	2979	61.4	572	14	US-10-216-159A-42	Sequence 42, Appl
22	2979	61.4	572	14	US-10-218-849-42	Sequence 42, Appl
23	2979	61.4	572	14	US-10-227-873-42	Sequence 42, Appl
24	2979	61.4	572	14	US-10-227-883-42	Sequence 42, Appl
25	2979	61.4	572	14	US-10-219-076-42	Sequence 42, Appl
26	2979	61.4	572	14	US-10-230-434-42	Sequence 42, Appl
27	2979	61.4	572	14	US-10-219-003-42	Sequence 42, Appl
28	2979	61.4	572	14	US-10-219-075-42	Sequence 42, Appl
29	2979	61.4	572	14	US-10-219-464-42	Sequence 42, Appl
30	2979	61.4	572	14	US-10-219-466-42	Sequence 42, Appl
31	2979	61.4	572	14	US-10-219-479-42	Sequence 42, Appl
32	2979	61.4	572	14	US-10-219-481-42	Sequence 42, Appl
33	2979	61.4	572	14	US-10-230-260-42	Sequence 42, Appl
34	2979	61.4	572	14	US-10-232-231-42	Sequence 42, Appl
35	2979	61.4	572	14	US-10-232-233-42	Sequence 42, Appl
36	2979	61.4	572	14	US-10-216-165-42	Sequence 42, Appl
37	2979	61.4	572	14	US-10-218-956-42	Sequence 42, Appl
38	2979	61.4	572	14	US-10-219-468-42	Sequence 42, Appl
39	2979	61.4	572	14	US-10-219-478-42	Sequence 42, Appl
40	2979	61.4	572	14	US-10-219-536-42	Sequence 42, Appl
41	2979	61.4	572	14	US-10-233-205-42	Sequence 42, Appl
42	2979	61.4	572	14	US-10-219-072-42	Sequence 42, Appl
43	2979	61.4	572	14	US-10-219-470-42	Sequence 42, Appl
44	2979	61.4	572	14	US-10-219-474-42	Sequence 42, Appl
45	2979	61.4	572	14	US-10-219-524-42	Sequence 42, Appl

ALIGNMENTS

RESULT 1  
US-09-735-705-225  
; Sequence 225, Application US/09735705  
; Patent No. US20020052329A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy R.  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Fanger, Neil  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C14  
; CURRENT APPLICATION NUMBER: US/09/735,705  
; CURRENT FILING DATE: 2000-12-12  
; NUMBER OF SEQ ID NOS: 419  
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 225  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-735-705-225

Alignment Scores:  
Pred. No.: 7,236-270 Length: 560  
Score: 2999.00 Matches: 559  
Percent Similarity: 99.82% Conservative: 0  
Best Local Similarity: 99.82% Mismatches: 1  
Query Match: 61.84% Indels: 1  
Gaps: 0

US-10-039-272-1 (1-2661) x US-09-735-705-225 (1-560)

```
QY 60 ATGGAATGCTCTACTATTCTCGGATTTCTGCTCTCGGCTGCAAGATTGCCATTGAT 119
Db 1 MetGluCysLeuTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCCGCCAAACGATTTCATGATGTGCTGGCAATGAAGACCTTCTGCTTACATGAGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAATGCGCTGCTCTGATGAAATGACTGGAATGAAAACTCTACCCA 239
Db 41 HisAsnGlnLeuAsnGlyTyrSerSerAspGluAsnAspTyrPheGluLeuTyrPro 60
QY 240 GTGTGGAAGCGGGGACATGAGGTGGAAAACTCTCGAAGGGAGCGCTGTGCAGCG 299
Db 61 ValTyrLysArgGlyAspMetArgTyrPheLeuLeuAlaAlaArgValGlnAla 80
QY 300 GTCTGACAGTACATCAGCCCTCGTGGCTCAAAATATACATTTCCGGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
QY 360 ATATTCCCTAGATGCCAAAGGAAGATGCCAATAGTCCAACTATGAGAAAGAACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAAATGAGCGTGGTTATCTGCTGATCCATATGTTTACAACCTGGACAGATGGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTyrThrAlaTyrSerGlu 140
QY 480 GACAGTACGGGGAAAATGCAACCGCCGCAAGCCATCAACCTCTCCCTGATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
QY 540 CCTTTTCTCACCACCGGATGGAGAAGATGGAATTTCACTACGCTCTTCCACACACTT 599
Db 161 ProPheProHisHisProGlyTyrArgArgTyrPhePheIleTyrValPheHisThrLeu 180
QY 600 GGTCAAGTATTCAGAAAATGGACAGATGTTTCAGTGAGATTTCTGTGAACACAGCCAA 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTACACATCTGGCTCACTCATGGAAGTCACTGTCTACAGAACATGCAAGCGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCCACAAGTGAAGATGTGCTGTTGTAACAGATCAGATTCTCTGTTGTTGTG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAAGACGATCGAAATTCATCCGACGAAACCTTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTTGTATCTCGATTTCATGATCCTAGCCACTTCTCAATATTCTACCATTAAC 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTAGTGGCTGTTTGTTCACCAATCATCTGTGAAT 958
```

```
Db 281 TyrLysTyrSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTGCTCAATGGAACCTTCAGCTTAACTCACTGTGAAAAGCTGCACACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTGTCCGCCACCCACCCACCCACCCAGACTTCAAAACCCACCCCTCTTTAGGA 1078
Db 321 GlyProCysProProProProProProProProProProSerLysProThrProSerLeuGly 340
QY 1079 CCTGCTGTGACAACCCCTCGAGCTGAGTGGATTCCTGATGAAACTGCCAGATTAAAC 1138
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
QY 1139 AGATATGCCACTTTCAGCCACCATCACAAATGTAGAGGGATCTTACAGATTAAACATC 1198
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
QY 1199 ATCCAGATGACAGACGCTCTGATGCCGTGCGATCGCTGAAAGCTCCCTAATAGACTTT 1258
Db 381 IleGlnMetThrAspValLeuMetProValProTyrProGluSerSerLeuIleAspPhe 400
QY 1259 GTCTGACCTGCCAAGGAGGACATTCACGAGAGTCTGTACCATCATTTCTGACCCACC 1318
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420
QY 1319 TCCGAGATCACCACAGACACAGCTCTGCGCTGCGATGTGGATGTGGATGAGATGTCTGCTG 1378
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY 1379 ACTGTGAGACGAACTTCAATGGGTCTGGAGCTACTGTGTGAACCTCACCTCGGGGAT 1438
Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460
QY 1439 GACACAAGCTCGCTCTACGAGACCCCTGATTTCTGTTCTCTGACAGACCCAGCTCG 1498
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
QY 1499 CCTTTAAGGATGCAAAACAGTCCCTGATCTCCGTTGGCTGTGTCCTATTTGTCACT 1558
Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY 1559 GTGATCTCCCTCTTGTGTGTACAAAACACAGGAATACAAACCAATAGAAAATAGTCCT 1618
Db 501 ValIleSerLeuLeuValTyrLysHisLysGluTyrAsnProIleGluAsnSerPro 520
QY 1619 GGAATGTGCTCAGAACAAAGCCCTGAGTGTCTTCTCAACCGTCAAAAGCCGTCTC 1678
Db 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY 1679 TTCCCGGAAACCCAGGAAAGGATCCGCTACTCAAAACCAAGAAATTTAAAGGAGTTCT 1738
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

RESULT 2
US-09-850-716A-225
; Sequence 225, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
```

ORGANISM: Homo sapien  
US-09-850-716A-225

## Alignment Scores:

Pred. No.: 7,23e-270 Length: 560  
Score: 2999.00 Matches: 559  
Percent Similarity: 99.82% Conservative: 0  
Best Local Similarity: 99.82% Mismatches: 1  
Query Match: 61.84% Indels: 1  
DB: 9 Gaps: 0

US-10-039-272-1 (1-2661) x US-09-850-716A-225 (1-560)

```

QY 60 ATGGAATGCTCTACTATTCTGGCAATTTCTGCTCTGGCTGCGAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuLeuAlaAArgLeuProLeuLeuAsp 20
QY 120 GCCGCCAAACGATTTCATGATGTGCTGGCGCAATGAAGACCTTCTGTTCATACAGGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAAATCAATTAATGGCTGCTCTCTGATGAATAATGATCGAATGAAAACTCTACCCA 239
Db 41 HisAsnGlnLeuAsnGlyTyrSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60
QY 240 GTGTGGAAGCGGGGAGACATGAGGTGGAAAAAACCCTCGAAGGGAGCCGTGCGACGG 299
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla 80
QY 300 GTCCTGACCATGACTCACACGCCCTCGTGGCTCAATATAACATTTGCCGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
QY 360 ATATTCCTAGATCCCAAGAGATGCCAATGGCAACATAGTCTATGAGAAACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAATGAGGCTGCTTTCTGCTGATCCATATGTTTACACTGACGACGATGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTrpValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACAGTGGGGGAAATGGCCGCGCAAGCCATCAACGCTTTCCCTGTATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
QY 540 CCTTTCTCCACACCGCGATGAGAGATGGAATTTCACTACGCTCTCCACACACTT 599
Db 161 ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCAATATTCACAGAAATGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTGGGCTCAACTCATGGAAGTACTGTCTACAGAAACATGGACGGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCCCAAGTGAAGATGTACTGTGTATACAGATCAGATTCCTGTGTTGTG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAGAACGATCGAATTCATCCGAGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTGATGCTCTGATTCATGATTCATGACCTTCCTCAATTAATCTACATTAAAC 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGGATATACCTGGCTGTTGTTTCCACCATACTATGTGAAT 958
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300

```

```

QY 959 CACACGTATGTGCTCAATGGAACTTACGCTTAACCTTACCTGTGAAAGCTCAGACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCCGCCACCGCCACCCACCCAGACCTTCAAAACCCACCCCTTCTTAGGA 1078
Db 321 GlyProCysProProProProProProProProProProSerLysProThrProSerLeuGly 340
QY 1079 CTTGCTGTGACAAACCCCTCGAGCTGAGTAGGATTCTCTGATGAAAACTGCCAGATTAAAC 1138
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
QY 1139 AGATATGCCCACTTTCAAGCCACCATCACATTCATAGGGGAATCTTAGAGTTAAATC 1198
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
QY 1199 ATCCAGATGACAGACGCTCTGATCCGCTGCCATCGGCTGAAAGCTCCCTAATAGACTTT 1258
Db 381 IleGlnMetThrAspValLeuMetProValProIleProIleProIleProIleProIle 400
QY 1259 GTCTGTGACCTGCCAAGGAGCATTCACACGAGGCTCTGTACCATCATTTCTGACCCACC 1318
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420
QY 1319 TGCAGATCACCCAGAACACAGCTCTGAGCCCTGTGATGATGATGATGATGATGATGATG 1378
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY 1379 ACTGTGAGACGAACTTCAATGGTCTGGGACGCTACTGTGACCTCACCTCCCTGGGGAT 1438
Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460
QY 1439 GACAAAGCCTGGCTCTCACGAGCACCTGATTTCTTCTCTGACAGACCCAGCCTCG 1498
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
QY 1499 CCTTTAAGATGGCAACAGCTGCCCTGATCCGCTGGCTGCTGGCCATATTTGTCACT 1558
Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY 1559 GTGATCTCCCTCTTGGTGTACAAAAACACAAAGAAATACAAACCAATAGAAATAGTCT 1618
Db 501 ValIleSerLeuLeuValTyrLysLysHisLysLysLysLysLysLysLysLysLys 520
QY 1619 GGAATGTGTCAGAAAGCAAGGCTGAGTGTCTTTCTCAACCGTGCAGAAAGCGCTGTC 1678
Db 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY 1679 TTCCCGGAAACAGGAAAGGATCCGCTACTCAAAACCAAGATTTAAAGGAGTTTCT 1738
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

```

## RESULT 3

US-09-897-778-225  
; Sequence 225, Application US/09897778  
; Patent No. US20020147143A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Warnerakis, Margarita  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedwick, Thomas S.  
; APPLICANT: Carter, Darick  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Peckham, David W.  
; APPLICANT: Fanger, Neil  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C16  
; CURRENT APPLICATION NUMBER: US/09/897,778  
; CURRENT FILING DATE: 2001-06-28  
; NUMBER OF SEQ ID NOS: 467  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 225  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-897-778-225

## Alignment Scores:

Pred. No.: 7,236-270 Length: 560  
Score: 2999.00 Matches: 559  
Percent Similarity: 99.82% Conservative: 0  
Best Local Similarity: 99.82% Mismatches: 1  
Query Match: 61.84% Indels: 1  
DB: 9 Gaps: 0

US-10-039-272-1 (1-2661) x US-09-897-778-225 (1-560)

```
QY 60 ATGGAATGCTCTACTATTTCTGGGATTTCTGCTCTGGCTGCAAGATTGCCATTGAT 119
DB 1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCGGCCAAACGATTTCATGATGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGAG 179
DB 21 AlaAlaIysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAAATCAATTAATGGCTGCTTCTGATGAAATGACTGGAATGAAATACTTACCCA 239
DB 41 HisAsnGlnLeuAsnGlyTyrSerSerAspGluAsnAspTyrPheAlaValAsnLeu 60
QY 240 GTGTGGAAAGCGGGAGACATGAGTGGAAACTCTCTGGAAGGAGCGGTGTGAGCGG 299
DB 61 ValTyrPheAspGlyAspMetArgTyrPheAsnSerTyrPheValGlyArgValGlnAla 80
QY 300 GTCCTGACAGTACTACAGCCCTCTGGGCTCAATATTAACATTTGCGGTGACCTG 359
DB 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
QY 360 ATATTCCCTAGATGCCAAAGGAAGATGCCAATGCAATAGTCTATGAGAAGAACTGC 419
DB 101 IlePheProArgCysGlnGlyGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAAATGAGGCTGTTTATCTGCTGATCCATATGTTTACAATGGACACATGTCAGAG 479
DB 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTyrThrAlaTyrSerGlu 140
QY 480 GACAGTACGGGAAATGCGACCGCCGAAAGCCATCAATACGCTCTCCCTCATGCGGAA 539
DB 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
QY 540 CCTTTTCTCACACCGCGATGGAGAAGATGGAATTTCTATCTACGCTCTTCCACACACTT 599
DB 161 ProPheProHisHisProGlyTyrArgArgTyrPheValPheIleTyrValPheHisThrLeu 180
QY 600 GGTCAATATTCAGAAATGGACGATGTCAGTGAGAGTTTCTGTGAACACAGCCAAT 659
DB 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTGGGCTCACTCATGGAAGTCACTGTCTACAGAGACATGAGCGGCATAT 719
DB 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCCCAAGTGAAGATGTACGTGGTAACAGATCAGATTCCTGTGTTGTG 779
DB 221 ValProIleAlaGlnValAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAAGACGATGAAATTCATCCGAGAAACCTTCCC-AAAGATCTCCCC 838
DB 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTTCATGCTGATTCATGATCCTAGCCACTTCTCAATATTCTACCATTAAC 898
DB 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTCGGGGATAATACTGGGCTGTTTGTTCACCAATCATCTGTGAAT 958
```

```
DB 281 TyrLysTyrSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAGCTGCACACCA 1018
DB 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTTCGCGCACCGCCACACACACACAGACCTTCAAAACCCACCCCTCTTTAGGA 1078
DB 321 GlyProCysProProProProProProProProProProProProProProSerLeuGly 340
QY 1079 CCTGCTGTGTAACCCCTCGAGCTGAGTAGGATTCTGATGAAACTGCCAGATTAAAC 1138
DB 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
QY 1139 AGATATGGCCACTTTCAAGCCACCATCAATTTAGAGGGAATCTTAGAGGTTTAACATC 1198
DB 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
QY 1199 ATCCAGATGACAGACGCTCTGATGCGGTGCCATGGCTGAAAGCTCCCTAATGACTTT 1258
DB 381 IleGlnMetThrAspValLeuMetProValProTyrProGluSerSerLeuIleAspPhe 400
QY 1259 GTCGTGACCTGCCAAGGAGGATTCACGAGGCTCTACCATCATTTCTGACCCACC 1318
DB 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420
QY 1319 TGCAGATGACCCAGACACAGCTCTGACGCCCTGTGGATGTGGATGTGGATGTGTCTGCTG 1378
DB 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY 1379 ACTGTGAGACGAACTTCAATGGGTCTGGAGCTACTGTGTGAACCTACCCCTGGGGAT 1438
DB 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460
QY 1439 GACAAAGCTGGCTCTCACGAGCCCTGATTTCTGTTCCTGACAGACACCCAGCTCG 1498
DB 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
QY 1499 CCTTTAAGGATGGCAACAGCTGCTGATCTCCGTTGGCTGTGGCTATTTGTCACT 1558
DB 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY 1559 GTGATCTCCTCTTGGTGTACAAAAACAACAGGAATACACCCAAATAGAAATAGTCTCT 1618
DB 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520
QY 1619 GGAATGTGTCAGAAAGCAAGCGCTGATGTTCTTCAACCGTGCAAAAGCCGTTTC 1678
DB 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY 1679 TTCCCGGAAACACAGAAAGGATCGCTACTCAAAACCAAGAAATTTAAAGAGTTTCT 1738
DB 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560
```

## RESULT 4

US-09-943-075A-6  
; Sequence 6, Application US/09943075A  
; Patent No. US20020151486A1  
; GENERAL INFORMATION:  
; APPLICANT: Popoff, Steven N.  
; APPLICANT: Safado, Faye F.  
; APPLICANT: Owen, Thomas A.  
; APPLICANT: Smock, Steven L.  
; TITLE OF INVENTION: Osteocalcin Protein and Nucleic Acids Encoding the Same,  
; FILE REFERENCE: 71369.262  
; CURRENT APPLICATION NUMBER: US/09/943,075A  
; PRIOR FILING DATE: 2001-08-30  
; PRIOR FILING DATE: 2000-08-30  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 4.0

! SEQ ID NO 6  
! LENGTH: 560  
! TYPE: PRT  
! ORGANISM: Human  
US-09-943-075A-6

## Alignment Scores:

Pred. No.: 7,23e-270 Length: 560  
Score: 2999.00 Matches: 559  
Percent Similarity: 99.82% Conservatives: 0  
Best Local Similarity: 99.82% Mismatches: 1  
Query Match: 61.84% Indels: 1  
DB: 9 Gaps: 0

US-10-039-272-1 (1-2661) x US-09-943-075A-6 (1-560)

```
QY 60 ATGGAATGCTCTACTATTCTCGGATTTCTGCTCGTGGCTGCAAGATTGCACTTGAT 119
Db 1 MetGluCysLeuTyrTyPheLeuGlyPheLeuLeuLeuAlaAArgLeuProLeuAsp 20
QY 120 GCGGCCAAACGATTTCATGATGCTGCTGGCAATGAAAGACCTTCGTCTCATGAGGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaLysMetArgGlu 40
QY 180 CACAATCAATTAATGCTGCTCTCTGATGAAATGACTGGAATGAAATACTCTACCCA 239
Db 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60
QY 240 GTGTGGAAAGCGGGAGACATGAGTGGGAAAACCTCTCGAAGGGAGGCGGTGTGAGGCG 299
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla 80
QY 300 GTCCTGACCGACTGACTCACAGCCCTCGTGGCTCAAAATATAAATTTCGGGTGAACCTG 359
Db 81 ValleuThrSerAspSerProAlaLeuValGlySerAsnLleThrPheAlaValAsnLeu 100
QY 360 ATATTCCCTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAAGACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnLleValTyrGluLysAsnCys 120
QY 420 AGAATGAGCTGGTTATCTGCTGATCCATATGTTTACAACTGGACAGCATGTCAGAG 479
Db 121 ArgAsnGlnAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACAGTCACGGGGAATGGACCGCCAAAGCCATCATACGTCCTCCCTGATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyLysSerHisAsnValPheProAspGlyLys 160
QY 540 CTTTTCCTCACCACCCCGGATGGAGAAGATGGGAATTTCTATCTACGCTTCCACACACTT 599
Db 161 ProPheProHisAsnProGlyTrpArgArgTrpAsnPhelleTyrValPheHisThrLeu 180
QY 600 GGTCAAGTATTCAGAAATGGGACGATGTCTAGTGAGAGTTTCTGTGAACACAGCCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGCGCTCAACTCAGTGAAGTCTGCTACAGAAGACATGGACGGGCAAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAAAGTGAAGATGTGTACGTGTGAAGATGGGAATTTCTCTGTTTGTG 779
Db 221 ValProileAlaGlnValLysAspValTyrValValThrAspGlnLleProValPheVal 240
QY 780 ACTATGTTCCAGAAAGACGATCGAAATTCATCCGAGGAAACCTTCCC-AAAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTGATCTCTGATTCATGATCTAGGCACTTCTCTCAATATTCTACCATTAAC 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCCCTGTTGTTTCCACCAATCATCTGTGAAT 958
```

```
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACAGCTATGTGCTCAATGGAACCTTCAGCCTTAACCTCAGTGTGAAGCTCAGCACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCCGCCACCGCCACCCACCCAGACCTTCAAAACCCACCCCTCTTTAGGA 1078
Db 321 GlyProCysProProProProProProProProProProProProProProProPro 340
QY 1079 CCTGCTGTGATCAACCCCTCGAGCTGATGAGTCTCTGATGATAAACTGCCAGATTAA 1138
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgLleProAspGluAsnCysGlnIleAsn 360
QY 1139 AGATATGCCCACTTTCAAGCCACCATCAATTTAGAGGGAATCTTAGAGTTAAATC 1198
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
QY 1199 ATCCAGATGACAGACGCTCCTGATGCGGTGCATGGCTGAAAGCTCCCTAATAGACTTT 1258
Db 381 IleGlnMetThrAspValLeuMetProValProTyrProGluSerSerLeuIleAspPhe 400
QY 1259 GTCTGACTCCCAAGGGAGCATTCACACGAGGCTGTGTACCATCATTTCTGACCCACC 1318
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420
QY 1319 TGCCAGATCACCCAGAACACAGTCTGAGCCCTGAGTGGATGGATGAGATGTCTGCTG 1378
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY 1379 ACTGTGAGAGCAACCTTCAATGGTCTGGGACGACTGTGTGAACCTCACCTCGGGGAT 1438
Db 441 ThrValArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460
QY 1439 GACACAGCTGGCTCTCACGAGCACCTGATTTCTGTTCTCTGACAGAGACCCAGCTCG 1498
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
QY 1499 CTTTAAAGATGGCAACAGTCCCTGATCTCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1558
Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY 1559 GTGATCTCCCTCTGTGTGTACAAAAACACAAAGGAATACAAACCAATAGATAAGTCT 1618
Db 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProLleGluAsnSerPro 520
QY 1619 GCGAATGTGTGAGCAAGGCTGAGTGTCTTTCTCAACCGTGCAGGCGGTGTTCT 1678
Db 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY 1679 TTCCCGGAAACCCAGGAAAGGATCCGCTACTCAAAACCCAGAAATTTAAGGAGTTTCT 1738
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560
```

## RESULT 5

US-10-007-700-225  
; Sequence 225, Application US/10007700  
; Publication No. US2003006497A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Fanger, Neil  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Durham, Margarita  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedwick, Thomas S.  
; APPLICANT: Carter, Darrick

APPLICANT: Watanabe, Yoshihiro  
APPLICANT: Peckman, David W.  
APPLICANT: Cai, Feng  
APPLICANT: Foy, Teresa M.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
FILE REFERENCE: 210121.455C17  
CURRENT APPLICATION NUMBER: US/10/007,700  
CURRENT FILING DATE: 2001-11-30  
NUMBER OF SEQ ID NOS: 469  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 225  
LENGTH: 560  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-007-700-225

Alignment Scores:  
Pred. No.: 7, 23e-270 Length: 560  
Score: 2999.00 Matches: 559  
Percent Similarity: 99.82% Conservative: 0  
Best local Similarity: 99.82% Mismatches: 1  
Query Match: 61.84% Indels: 1  
DB: 12 Gaps: 0

US-10-039-272-1 (1-2661) x US-10-007-700-225 (1-560)

QY 60 ATGGAAATGCTCTACTATTCTCGGATTCTCGCTCGCTCGCTCGCAAGATTGCCACTGAT 119  
DB 1 MetGluCysLeuTyrPheLeuGlyPheLeuLeuLeuAlaAlaArgLeuProLeuAsp 20  
QY 120 GCGCCAAACGATTTCATGATGCTGGCAATGAAGACCTTCGTTACATGAGGAG 179  
DB 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40  
QY 180 CACAATCAATTAATGGCTGCTTCTGATGAAATGACTGGAATGAAATCACTACCCA 239  
DB 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTyrAsnGluLysLeuTyrPro 60  
QY 240 GTGTGAACGGGGAGACATGAGTGGAAAACTCTCGAAGGGGCGGTGTCAGGCG 299  
DB 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla 80  
QY 300 GTCCTGACAGTCACTACAGCCCTCGTGGCTCAATATAACATTTGCGGTGAACCTG 359  
DB 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnLeuThrPheAlaValAsnLeu 100  
QY 360 ATATTCCTTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGACTGC 419  
DB 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120  
QY 420 AGAATGAGCGTGTATCTGCTGATTCATATGTTTACAACTGGACAGCATGGTCAGAG 479  
DB 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140  
QY 480 GACAGTACGGGGAATGCAACCGCCAAAGCCATCAATCACTCTTCCTGATGGGAAA 539  
DB 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisHisAsnValPheProAspGlyLys 160  
QY 540 CTTTTTCTCACCACCCGAGATGGAGAAGATGGAATTTCTATCTACGTTCTCCACACTT 599  
DB 161 ProPheProHisHisProGlyTrpArgTrpAsnPheIleTyrValPheHisThrLeu 180  
QY 600 GGTCAATTTCCAGAAATGGACGATGTTTCAGTGAGAGTTCTCTGACACAGCAAT 659  
DB 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200  
QY 660 GTGCACTTGGGCTCAACTCATGGAAGTCACTGTCTACAGAAGACATGGACGGCATAT 719  
DB 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220  
QY 720 GTTCCCATCGCACAAGTGAAGATGTGTACGTGTGAACAGATCTCTGTGTGTGTGT 779

DB 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240  
QY 780 ACTATGTTCCAGAAACGATCGAAATTCATCCGACGAAACCTTTCCC-AAAGATCTCCC 838  
DB 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260  
QY 839 ATTATGTTGATGCTGATTTCATGATTCATGATTCATGATTCATGATTCATGATTCAT 898  
DB 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280  
QY 899 TACAAGTAGGCTCGGGGATAATACGGCTGTTGTTTCCACCAATCATCTCTGAT 958  
DB 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300  
QY 959 CACACGTATGTGCTCAATGGAACCTTCAGCCCTTAACCTCCTGTAAGAGTCGACACCA 1018  
DB 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaAlaPro 320  
QY 1019 GGACCTTGTCCGACCGCCACCCACCCAGACCTTCAAAACCCACCCCTCTTTAGGA 1078  
DB 321 GlyProCysProProProProProProProProProProSerLysProThrProSerLeuGly 340  
QY 1079 CCTGCTGTGTGACACCCCTCGAGCTGAGTAGGATTCCTGATGAAAACTGCCAGATTAAAC 1138  
DB 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360  
QY 1139 AGATATGGCCACTTTCAAGCCACCATCAATATTGAGAGGGAATCTTGAAGTTAAATC 1198  
DB 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380  
QY 1199 ATCCAGATGACAGACGCTCCTGATGCGGTGCGATCGCTGAAGCTCCCTAATAGACTTT 1258  
DB 381 IleGlnMetThrAspValLeuMetProValProTyrProGluSerSerLeuIleAspPhe 400  
QY 1259 GTCGTGACCTGCAAGGGAGCATTCGCCGAGGTCTGTACCATCATTTCTGACCCCAAC 1318  
DB 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420  
QY 1319 TCGAGATCACCCAGAACACACAGCTCTCCAGCCCTGCGATGTGGATGTGGATGTGTGCTG 1378  
DB 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440  
QY 1379 ACTGTGAGACGAACTTCAATGGTCTGGAGCTGCTGTAACCTCACTACCTACCTGGGGAT 1438  
DB 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460  
QY 1439 GACACAGCTCGCTCTCACGAGACCCCTGATTTCTGTTTCTGACAGACACCCAGCTCG 1498  
DB 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480  
QY 1499 CCTTTAAGGATGGCAACAGTGCCTGATCTCCGTTGGTGTGCTGGCCATATTTGTCACT 1558  
DB 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500  
QY 1559 GTGATCTCCTCTGTGTGTACAAAAACACAGGAATACAAACCAATAGAAATAGTCCT 1618  
DB 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520  
QY 1619 GCGAATGTGTGAGAGCAAGAGCCCTGAGTGTCTTCTCAACCGTCAAAAGCCGTGTTTC 1678  
DB 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540  
QY 1679 TTCCCGGGAACACGAGAAAGGATCCGCTACTCAAAACCAAGATTTAAAGAGATTCT 1738  
DB 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

## RESULT 6

US-10-463-106-2  
; Sequence 2, Application US/10463106  
; Publication No. US20030202938A1  
; GENERAL INFORMATION:  
; APPLICANT: RAMESHWAR, Pranela  
; TITLE OF INVENTION: HEMATOPOIETIC GROWTH FACTOR INDUCIBLE NEUROKININ-1 GENE

```

, FILE REFERENCE: 267/033 University of Medicine & Dentistry of New Jersey
, CURRENT APPLICATION NUMBER: US/10/463.106
, CURRENT FILING DATE: 2003-06-17
, PRIOR APPLICATION NUMBER: US 10/039,272
, PRIOR FILING DATE: 2001-10-20
, PRIOR APPLICATION NUMBER: US 60/241,881
, PRIOR FILING DATE: 2000-10-20
, NUMBER OF SEQ ID NOS: 2
, SOFTWARE: PatentIn version 3.1
, SEQ ID NO 2
, LENGTH: 560
, TYPE: PRT
, ORGANISM: Homo sapiens
US-10-463-106-2

```

Alignment Scores:				
Pred. No.:	7,23e-270	Length:	560	
Score:	2999.00	Matches:	559	
Percent Similarity:	99.82%	Conservative:	0	
Best Local Similarity:	99.82%	Mismatches:	1	
Query Match:	61.84%	Indels:	1	
DB:	12	Gaps:	0	
US-10-039-272-1 (1-2661) x US-10-463-106-2 (1-560)				
QY	60	ATGGAATGTCCTACTATTTCCTGGATTTCCTGCTCCTGGCTCAAGATTGCCACTTGAT	119	
Db	1	MetGluCysLeuTyrPheLeuGlyPheLeuLeuLeuAlaAlaArgLeuProLeuAsp	20	
QY	120	GCGGCCAAACGATTCATGATGTCTGGCCAAATGAAAGACCTTCGTCTACATCAGGGAG	179	
Db	21	AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu	40	
QY	180	CACAATCAATTAAATGGCTGCTCTTCATGAAATGACTGGAAATGAAAACTCTACCA	239	
Db	41	HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro	60	
QY	240	GTGTGGAAGCGGGAGACATGAGTGGAAAAAATCTCTGGAAGGAGGCGGTGTCAGGCG	299	
Db	61	ValTTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla	80	
QY	300	GTCTCGACCAAGTACTACAGCCCTCGTGGGTCAAAATATACATTTGGCGTGAACCTG	359	
Db	81	ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu	100	
QY	360	ATATTCCCTAGATGCCAAAAGAAAGATGCCAATGGCAACATAGTCTATGAGAGAACTCG	419	
Db	101	IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys	120	
QY	420	AGAAATGAGCTGGTTTATCTGCTGATCCATATGTTTACAACTGGACAGCATGGTCAGAG	479	
Db	121	ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu	140	
QY	480	GACAGTCACGGGAAAAATGGCACCGCAAGCCATCATAACTCTTCCTCATGCGGAA	539	
Db	141	AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys	160	
QY	540	CCTTTTCTCACCACCCGGATGGAGAAGATGGAATTCATCTACGCTCTCCACACACTT	599	
Db	161	ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu	180	
QY	600	GTCTAGTATTTCCAGAATTTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACACGCAAT	659	
Db	181	GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn	200	
QY	660	GTGACACTTGGGCCTCAACTCATGGAAGTGAAGTCTTACAGAAGACATGGACGGGCATAT	719	
Db	201	ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr	220	
QY	720	GTTCCTCATCGACAGTGAAGATGTGTACGTGGTAAACATCAGATTCTCTGTGTTGTG	779	
Db	221	ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIlePrValPheVal	240	

QY	780	ACTATGTTCCAGAGAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAAGATCTCCCC	838
Db	241	ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro	260
QY	839	ATTATGTTGATGTCCTGATTCATGATCCTTAGCCACTTCCTCAATATTCTACCATTAAC	898
Db	261	IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn	280
QY	899	TACAAGTGGACCTTCGGGGATAAATACTGGCCCTGTTTGTTCACCAATCATACTGTGAAT	958
Db	281	TyrLysItrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn	300
QY	959	CACAGGTATGTCTCAATGGAACTTCAGGCTTAACTCCTACTGTGAAGCTGCACGCCA	1018
Db	301	HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaAspPro	320
QY	1019	GGACTTGTCCGCGCACCGCCACACACACAGACCTTCAAAACCCACCCCTTCTTTAGGA	1078
Db	321	GlyProCysProProProProProProArgProSerLysProThrProSerLeuGly	340
QY	1079	CCTGCTGGTGACAAACCCCTCGGAGCTGAGTAGGATTCCTGATGAAAACTGCACATTAAAC	1138
Db	341	ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn	360
QY	1139	AGATATGGCCACTTTCAGGCCACCATCACAAATCTAGAGGAACTTTAGAGGTAAACATC	1198
Db	361	ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle	380
QY	1199	ATCCAGATGACAGAGTCCTGATGCCGGTCCCATGGCTCAAGAGTCCCTAAATAGACTTT	1258
Db	381	IleGlnMetThrAspValLeuMetProValProItrpProGlnSerSerLeuIleAspPhe	400
QY	1259	GTCGTGACCTGCCAAGGAGCATTCCCACGGAGGCTCTGTACCATCATTTCTGACCCCAACC	1318
Db	401	ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr	420
QY	1319	TGGAGATACCCAGAAACACAGTCTGAGCCCTGTGGATGTGGATGAGTGTCTGCTGCTG	1378
Db	421	CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu	440
QY	1379	ACTGTGAGACGAACCTTCAATGGTCTGGGACGCTACTGTGTGAACCTCAACCTGGGGAT	1438
Db	441	ThrValArgArgThrPheAsnGlySerGlyThyTyrCysValAsnLeuThrLeuGlyAsp	460
QY	1439	GACACAAGCTGCTCTACGAGCACCCCTCAATTCTGTTCTCTGACAGACACCCAGCCTCG	1498
Db	461	AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer	480
QY	1499	CCTTTAAGGATGCAACACAGTGCCTGATCTCGTTGGCTGCTTGGCCATATTGTGCACT	1558
Db	481	ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr	500
QY	1559	GTGATCTCCCTCTGTGTGTACAAAAACAACAGGAATACAAACCAATAGAAATAGTCTCT	1618
Db	501	ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGlnAsnSerPro	520
QY	1619	GGGAATGTGGTCAGAGCAAAAGCCTGAGTGTCTTCTCAACCGTCAAAAGCCGCTGTTTC	1678
Db	521	GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe	540
QY	1679	TTCCCGGGAAACACAGAAAGGATCCGCTACTCAAAACCAACAGAAATTHAAGAGTTTCT	1738
Db	541	PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer	560

## RESULT 7

US-10-039-272-2  
; Sequence 2, Application US/10039272  
; Publication No. US20020168653A1  
; GENERAL INFORMATION:  
; APPLICANT: RAMESHWAR, Pranela  
; TITLE OF INVENTION: HEMATOPOIETIC GROWTH FACTOR INDUCIBLE NEUROKININ-TYPE  
; FILE REFERENCE: 267/033 University of Medicine & Dentistry of New Jersey  
; CURRENT APPLICATION NUMBER: US/10/039,272



; CURRENT FILING DATE: 2001-10-20  
; PRIOR APPLICATION NUMBER: US 60/241,881  
; PRIOR FILING DATE: 2000-10-20  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-039-272-2

## Alignment Scores:

Pred. No.: 7, 23e-270 Length: 560  
Score: 2999.00 Matches: 559  
Percent Similarity: 99.82% Conservatives: 0  
Best local Similarity: 99.82% Mismatches: 1  
Query Match: 61.84% Indels: 1  
DB: 13 Gaps: 0

US-10-039-272-1 (1-2661) x US-10-039-272-2 (1-560)

```
QY 60 ATGGAAATGCTCTACTATTCTCGGATTTCTGCTCCTGGCTGCAAGATTGCCACTGAT 119
Db 1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCGGCCAAACGATTTCATGATGCTGCTGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAATGCTGCTTCTCTGATGAAATGACTGGAATGAAAACTTCAACCA 239
Db 41 HisAsnGlnLeuAsnGlyTyrSerSerAspGluAsnAspTyrAsnGluLysLeuTyrPro 60
QY 240 GTGTGGAACCGGGAGACATAGAGTGGAAAAAATCTCTCTGGAAGGAGCGCTGTGAGGGG 299
Db 61 ValTyrLysArgGlyAspMetArgTyrLysAsnSerTyrLysGlyGlyArgValGlnAla 80
QY 300 GTCCTGACCACTGACTCACCAGCCCTCGTGGCTCAATATAACATTTCCGTTGACACTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnLeuThrPheAlaValAsnLeu 100
QY 360 ATATTCCTTAGATGCCAAAGGAAGATGCCAATGCAACATAGTCTTACAGAACTGCTC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnLeuValTyrGluLysAsnCys 120
QY 420 AGAATGAGCGTGTATCTGCTGATCCATATGTTTACAACTGGACAGCATGGTTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTyrThrAlaTyrSerGlu 140
QY 480 GACAGTGACGGGAAATGSCACCGCCAAAGCCATCAATACCTCTCCCTGATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
QY 540 CTTTTTCCTCACCACCCCGGATGGAGAAGATGGAATTTCACTACGCTTCCACACACTT 599
Db 161 ProPheProHisHisProGlyTyrArgArgTyrAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCAATATTTCCAGAAATTTGGACAGATGTTTCAGTGAGATTTCTGTGAACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTGGGCTCAACTCATGGAAGTCACTGTCTACAGAAGACATGGACGGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCCTGTGTGTGTG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnLeuProValPheVal 240
QY 780 ACTATGTTCCAGAAAGACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
```

## RESULT 8

US-10-117-982-225

; Sequence 225, Application US/10117982

; Publication No. US20030138438A1

; GENERAL INFORMATION:

; APPLICANT: Foy, Teresa M.

; APPLICANT: Fanger, Gary R.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Henderson, Robert A.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Mericlie, Barbara

```
QY 839 ATTATGTTGATGCTCGATTTCATGATCCTAGCCACCTTCTCAATATTCTACCATTAAC 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGATAAATACGCGCTGTTGTTTCCACCAATCATACTGTGAAT 958
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTGCTCAATGGAAACCTTCAGCCCTTAACTCTACTGTGAAAGCTGCACACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCCGCCACCGCCACCCACCCAGACTTCAAACCCACCCCTTCITTAGA 1078
Db 321 GlyProCysProProProProProProProProProProProProProProProProPro 340
QY 1079 CTTGCTGTGTGACAAACCCCTCGAGCTGAGTAGGATTCCTGATGAAACCTGCCAGATTAAAC 1138
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
QY 1139 AGATATGCCCACTTTCAAGCCACCATCACAAATTTGAGAGGGAATCTTAGAGTTTAAATC 1198
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
QY 1199 ATCCAGATGACAGAGCTCCTGATGCCGCTGCGCTGAGAGCTCCCTAATAGACTTT 1258
Db 381 IleGlnMetThrAspValLeuMetProValProThrProGluSerSerLeuIleAspPhe 400
QY 1259 GTCGTGACTGCTCCAAAGGAGGACTTCCACGGAGGTCTGTACCATCATTTCTGACCCACCC 1318
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420
QY 1319 TCCGAGATGACCCAGAACACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTCTGCTG 1378
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY 1379 ACTGTGACAGCAACTTCAATGGGTCTGGGAGCTGAGTGTGAACTCACCTCACCCTGGGGAT 1438
Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460
QY 1439 GACACAGCTGGCTCTCACAGACCCCTGATTTCTGTTCTCTGACAGACCCAGCTCG 1498
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
QY 1499 CTTTTAAGAGATGGCAAAACAGTGCCTGATCTCCGTTGGCTGCTGCTGCCATATTTGTCACT 1558
Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY 1559 GTGATCTCCTCTTGTGTGTACAAAAACACAGGAATACACCCATAGAAATAGTCCT 1618
Db 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520
QY 1619 GGAATGTGCTCAGAGCAAGCCCTGAGTGTCTTTCTCAACCGTGCAGAACGCCGTGTC 1678
Db 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY 1679 TTCCCCGGGAAACCCAGGAAAGGATCCGCTACTCTCAAAAAACAAGATTTAAAGAGTTTCT 1738
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560
```



```

/ APPLICANT: Spies, Gregory A.
/ APPLICANT: Fan, Liqun
/ APPLICANT: Wang, Tongtong
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.455C18
/ CURRENT APPLICATION NUMBER: US/10/117,982
/ CURRENT FILING DATE: 2002-04-05
/ NUMBER OF SEQ ID NOS: 484
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 225
/ LENGTH: 560
/ TYPE: prt
/ ORGANISM: Homo sapiens
/ US-10-117-982-225

```

Alignment Scores:		
Pred. No.:	7,23e-270	Length:
Score:	2999.00	Matches:
Percent Similarity:	99.82%	Conservative:
Best Local Similarity:	99.82%	Mismatches:
Query Match:	61.84%	Indels:
DB:	14	Gaps:
		0
		0
		560
		559

US-10-039-272-1 (1-2661) x US-10-117-982-225 (1-560)

60	ATGGAATGCTCTACTATTTCCTGGGAAATTCCTGCTCTCGCTGCAAGATTGCACACTGAT	119
1	MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp	20
120	GCGCCCAACGATTTCATGATGCTGGCGCAATGAAGACCTTCGCTTACATGAGGAG	179
21	AlaAlaIysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu	40
180	CACAATCAATTAATGGCTGCTCTCTGATGGAATGACTGGAAAGAAAACCTTACCCA	239
41	HisAsnGlnLeuAsnGlyTyrPseSerAspGluAsnAspTrpAsnGluLysLeuTyrPro	60
240	GTGTGGAAGCGGGAGACATGAGGTGGAATAAATCCTCGAAGGAGGCGCTGTGCAGCG	299
61	ValTrrpIysArgGlyAspMetArgTrrpIysAsnSerTrrpIysGlyGlyArgValGlnAla	80
300	GTCTGTGACCACTGACTCACACGCCCTGTGGCTCAATATAACATTTTCGGTGAACCTG	359
81	ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu	100
360	ATATTCCCTAGATCCCAAAAGGAAGATGCCAATGCCAATAGTCTATGAGAGNACTGC	419
101	IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys	120
420	AGAAATGAGCCTGGTTTATCTGTGATCCATATGTTTCAACTGGACAGCATGTCAGAG	479
121	ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu	140
480	GACAGTCACGGGAAATATGCACCGGCCAAAGCCATCAACGCTTCCCTGATGGGAAA	539
141	AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys	160
540	CTTTTTCCTCACACCCCGGATGAGAGAGATGGAATTCATCTACGCTTCCACACATT	599
161	ProPheProHisHisProGlyTrrpArgGrrpAsnPheIleTyrValPheHisThrLeu	180
600	GGTCAGTATTTCAGAAATTTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAT	659
181	GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn	200
660	GTACACATTGGGCTCACTCATCGAAGTGACTGTCTTACAGAAGCATGGACGGGCGATAT	719
201	ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr	220
720	GTTCCTCATCGCAAGTGAAGATGTGACTGGTAACTACAGATCAGATTCCTGTGTTGTG	779
221	ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal	240

## RESULT 9

US-10-295-027-1258  
 ; Sequence 1258, Application US/10295027  
 ; Publication No. US20030232350A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Afar, Daniel  
 ; APPLICANT: Aziz, Natasha  
 ; APPLICANT: Ginsberg, Wendy M.

```
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1258
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1258

Alignment Scores:
Pred. No.: 7,236-270 Length: 560
Score: 2999.00 Matches: 559
Percent Similarity: 99.82% Conservative: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 61.84% Indels: 1
DB: 15 Gaps: 0

US-10-039-272-1 (1-2661) x US-10-295-027-1258 (1-560)
QY 60 ATCGAATGCTCTACATATTTCCTGGGATTTCTGCTGCTGCGAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCGGCCAAAGATTGATGATGCTGGGCAATCAAGACCTCTGCTTACATGAGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAATGCTGCTCTCTGATGAAATGACTGGAATGAAAACTTACCCA 239
Db 41 HisAsnGlnLeuAsnGlyTyrSerSerAspGluAsnAspIrpAsnGlnLysLeuTyrPro 60
QY 240 GTGTGGAGCGGGGAGACATGAGGTGGAATAAATCTCTGGAGAGGAGCCGTGTGAGCG 299
Db 61 ValTyrLysArgGlyAspMetArgTyrLysAsnSerTrpLysGlyGlyArgValGlnAla 80
QY 300 GTCTGACCACTGACTCACAGCCCTCTGGGCTCAAAATATAACATTTGGGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
QY 360 ATATTCCCTAGATGCCAAAAGAGATGCCAATGGCAACATAGTCTTAGAGAAACTGC 419
```

```
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAAATGAGGCTGTTTATCTGCTGATCCATATGTTTACAACTGGACAGCATGGTCAAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTyrThrAlaTyrSerGlu 140
QY 480 GACAGTGACCGGGAAATGGCACCGGCAAGCCATATACAGCTTCTCCCTGATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
QY 540 CTTTCTCTCACCCCGGATGGAGAGATGGAAATTTCACTAGCTTCTCCACACACTT 599
Db 161 ProPheProHisHisProGlyTyrArgArgTyrPheIleTyrValPheHisThrLeu 180
QY 600 GGTCAGTATTTCCGAAATTCGACGATGTTTCAGTGAGAGATTTCTGTGAACACGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGGCTCAACTCATGGAAGTGTCTACAGAGCATGGACGGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAGTGAAGATGTACGTGGTAAACAGATCAGATTCCTCTGTGTGG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAGACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTGATGTCCTGATTCATGATCTAGCCACTTCTCCTCAATTAATCTACCATTAAC 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGATTAATCTGCCCTGTTGTTTCCACCAATCATACTGTGAAT 358
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACAGTATGTCCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCCGACCGCCACACACCCAGACCTTCAAAACCCACCCCTCTTTAGGA 1078
Db 321 GlyProCysProProProProProProProProProProProProProProProSerLeuGly 340
QY 1079 CCTGCTGGTGACAAACCCCTGGAGCTGAGTATCTCTGATGAAACCTGCCAGATTAAAC 1138
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
QY 1139 AGATATGGCCACTTTCAAGCCACCATCAACAATTTAGAGGGAATCTTAGAGGTTAAATC 1198
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
QY 1199 ATCCAGATGACAGACCTCTGATGCGCGTGCATGGCTGAAAGCTCCCTTAATAGACTTT 1258
Db 381 IleGlnMetThrAspValLeuMetProValProValProGluSerSerLeuIleAspPhe 400
QY 1259 GTGCTGACCTGCCAAGGGAGCATTCACCGAGGCTGTGTACCATCATTTCTGACCCCAAC 1318
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420
QY 1319 TCGAGATCACCCAGAACACAGTCTGACGCCCTGTGGATGTGGATGTGAGATGTCTGCTG 1378
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY 1379 ACTGTGAGACGAACCTTCAATGGGTCTGGACGCTGCTGTAACCTCACCTCCCTGGGGAT 1438
Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460
QY 1439 GACAGAGCTGCTCTCACGAGCACCTGATTTCTCTCTCTCACAGAGACCCAGCCTCG 1498
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
```

QY 1499 CCTTTAAGGATGCAACAGTGCCTGATCTCCGTTGGCTGCTGGCCATTTTGTCACT 1558  
Db 481 ProLeuArgMetAlaAsnSerAlaLeuLeSerValGlyCysLeuAlaIlePheValThr 500  
QY 1559 GTGATCTCCCTCTTGTGTACAAAAACACAGGAATCAACACCAATAGAAAATAGTCCT 1618  
Db 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520  
QY 1619 GCGAATGTGTCAGAGCAAGCCCTGAGTGTCTTCTCAACCGTCAAAAGCCGTGTTTC 1678  
Db 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540  
QY 1679 TTCCCGGGAACACAGAAAAGGATCCGCTACTCAAAAAACAGAAATTTAAAGAGTTTCT 1738  
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

## RESULT 10

US-10-313-986-225  
; Sequence 225, Application US/10313986  
; Publication No. US20030236209A1  
; GENERAL INFORMATION:  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: McNabb, Andria  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Wang, Tongtong  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C19  
; CURRENT APPLICATION NUMBER: US/10/313,986  
; CURRENT FILING DATE: 2002-12-04  
; NUMBER OF SEQ ID NOS: 560  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 225  
; LENGTH: 560  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-313-986-225

Alignment Scores:  
Pred. No.: 7,23e-270 Length: 560  
Score: 2999.00 Matches: 559  
Percent Similarity: 99.82% Conservative: 0  
Best Local Similarity: 99.82% Mismatches: 1  
Query Match: 61.84% Indels: 0  
DB: 15 Gaps: 0

US-10-039-272-1 (1-2661) x US-10-313-986-225 (1-560)

QY 60 ATGGAATGTCCTACTATTTCTGGGATTTCTGCTGCTGCTGCTGCAAGATTGCCATTGAT 119  
Db 1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20  
QY 120 GCCGCCAAACGATTTCATGATGTGCTGGCAATCAAGACCTTCTGCTTACATCAGGAG 179  
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40  
QY 180 CACAAATCAATTAATGCTGCTTCTGATGAAATGACTGGAATGAAAACTCTACCCA 239  
Db 41 HisAsnGlnLeuAsnGlyTyrSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60  
QY 240 GTGTGGAGCGGGAGACATGAGTGGCAAAACTCTGGAGGAGGCGGTGTCAGGCG 299  
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla 80  
QY 300 GTCTGTACCACTGACTCACAGACCTCTGCTGGCTCAAAATATACATTTGGGTCAACTG 359  
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100  
QY 360 ATATTCCCTAGATCCCAAAAGGAGATGCCAATGGCAACATAGTCTATGAGAGAACTGC 419  
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120

QY 420 AGAAATCAGCGTGGTTTATCTGCTGATCCATATGTTTCAACCTGACAGCATGTCAGAG 479  
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140  
QY 480 GACAGTACGGGAAAATGCGCACCGCCAAAGCCATCATACGCTCTCCCTGTATGGAAA 539  
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisHisAsnValPheProAspGlyLys 160  
QY 540 CTTTTTCTCACACCCCGGATGGAGAGATGGAATTTTCATCTGCTCTCCACACACTT 599  
Db 161 ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180  
QY 600 GTCAGTATTTCCAGAAATTTGGAGCATGTTTCAGTGGAGATTTCTGTGAACACAGCAAT 659  
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200  
QY 660 GTGACACTTGGSCCTCACTCATGGAAGTACTGTCTACAGAGACATGACGGGCATAT 719  
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220  
QY 720 GTTCCCATCGCAAAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCCTGTGTTGTG 779  
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240  
QY 780 ACTATGTTCCAGAGAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCC 838  
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260  
QY 839 ATTAATGTTGATGCTCTGATTCATGATCCTAGCCACTTCCTCAATTAATTCACCAATAC 898  
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280  
QY 899 TACAAGTGGAGCTTCGGGGTAATACTGCGCTGTTGTTTCCACCAATCATTCTGTGAAT 958  
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300  
QY 959 CACAGTATGTGCTCAATGGAACTTCAGCTTAACTCTGCTGCTGCTGCTGCTGCTGCTG 1018  
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaLysPro 320  
QY 1019 GGACTTGTGCTGCGCACCGCCACACCCAGACCTTCAAAACCCACCCCTCTTTAGGA 1078  
Db 321 GlyProCysProProProProProProProProProProProProProProProSerLeu 340  
QY 1079 CCTGCTGTGACAAACCCCTGAGCTGAGTGGATTCCTGATGAAACTGCCAGATTAAC 1138  
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360  
QY 1139 AGATATGCGCACTTTCAAGCCACCATCAATTCAGAGGAACTTTAGAGGTTAAACATC 1198  
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380  
QY 1199 ATCCAGATGACAGACGTCCTGATGCGGTGCATGGCTGAAAGCTCCCTAATAGACTTT 1258  
Db 381 IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe 400  
QY 1259 GTGCTGACCTGCCAAGGAGCATTCACCGAGGTCTGTACCATCATTTCTGACCCCAAC 1318  
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420  
QY 1319 TGGAGATCAACCCAGAACACAGTCTGCGCCCTGTGGATGTGATGATGATGTGCTGCTG 1378  
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440  
QY 1379 ACTGTGAGACGAACTTCAATGGGTCTGGACCTACTGTGTGAACCTCACCTCGGGGAT 1438  
Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460  
QY 1439 GACAAAGCTGCTCTCAGACGACCTGATTTCTGTCTCTCCTACAGAGACCCAGCTCG 1498  
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480



```
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCTGTTGTTTCCACCAATCATCTGCAAT 958
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTCTCAATGGAACTTCAGCTTAACTCTCACTGTGAAAGCTGCAGACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCCGCCACCGCCACCAACCAACAGACCTTCAAAACCCCTCTTTAGGA 1078
Db 321 GlyProCysProProProProProProProProProProProProProProSerLeuGly 340
QY 1079 CTGTCTGGTGACAAACCCCTGGAGCTGAGTAGATTCTTGATGAAACCTGCCAGATTAC 1138
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
QY 1139 AGATATGGCCACTTTCAGACCCCAATCAAAATGTAGAGGGAATCTTAGAGTTAAATC 1198
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyLeuGluValAsnIle 380
QY 1199 ATCCAGATCAGACAGCTGCTGATGCCGTGCCATGCCCTGAAAGCTCCCTAATAGACTTT 1258
Db 381 IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe 400
QY 1259 GTCGTGACCTGCCAAGGGAGCAATCCACGGAGGCTGTGACCATCATTTCTGACCCACC 1318
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleLeuSerAspProThr 420
QY 1319 TCCGAGATCACCAGAACACAGCTGCGACCTGTGGATGGATGAGATGTCTGTGCTG 1378
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY 1379 ACTGTGACAGAACTTCAATGGGCTGGGACGTAAGTGTGCAACTCACCCCTGGGGAT 1438
Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460
QY 1439 GACAAAGCTGGCTCTCAGCAGACCTGATTTCTGTTCTGTTCTGAGAGACCCAGCTCG 1498
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
QY 1499 CCTTAAGATGGCAACAGTGCCTGATCCGTTGGCTGCTGTTGCCATATTTGTCACT 1558
Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY 1559 GTGATCTCCCTCTGGTGTACAAAACACAGGAATACACCCCAATAGAAAATAGTCTCT 1618
Db 501 ValIleSerLeuLeuValTyrLysHisLysGluTyrAsnProIleGluAsnSerPro 520
QY 1619 GGGATGTGTGAGAACAAAGGCTGATCCGTTGGCTGCTGTTGCCATATTTGTCACT 1678
Db 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY 1679 TTCCCGGAACACAGGAAGGATCCGCTACTCAAAACCAAGAAATTAAGGAGTTTCT 1738
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560
```

## RESULT 12

```
US-10-408-765A-466
; Sequence 466, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
```

```
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 466
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-466
```

## Alignment Scores:

```
Pred. No.: 7,236-270 Length: 560
Score: 2399.00 Matches: 559
Percent Similarity: 99.82% Conservativeness: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 61.84% Indels: 1
DB: 16 Gaps: 0
```

US-10-039-272-1 (1-2661) x US-10-408-765A-466 (1-560)

```
QY 60 ATGGAATGCTCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCCGCCAAACGATTTTCATGATGTGTGGCAATGAAAGACCTTCTGCTTACATGAGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAATGCTGCTCTCTGATGAAATGACTGGAATGAAAACTCTACCCA 239
Db 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60
QY 240 GTGTGGAAGCGGGAGACATGAGGTGGAATAAATCTCTGGAAGGGAGCGCTGTGCAGCG 299
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla 80
QY 300 GTCTGTGACCTGACTCACCAGCCCTCGTGGCTCAAAATAATAATTTCCGGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
QY 360 ATATTCCCTAGATGCCAAAGAGATGCCAATGGCAACATAGTCTATGAGAACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAAATGAGGCTGTTTCTGCTGATCCATATGTTTACAACTGAGCAGCATGCTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnIleThrAlaTrpSerGlu 140
QY 480 GACAGTACCGGGGAAATGGCACCGCCAAAGCCATCATACCTCTTCCCTGTGAGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisHisAsnValPheProAspGlyLys 160
QY 540 CCTTTTCTCACACCCCGATGGAGAGATGGATTTTCATCTACGTCTTCCACACTT 599
Db 161 ProPheProHisHisProGlyTrpArgTrpAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCACTATTTCAGAAATTTGGACAGATGTTTCAGTGAGAGTCTTGTGGAACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTGGGCTCAACTCATGGAAGTACTGTCTACAGAAGACATGGACGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCAACAGTGAAGATGTACGTGGTAAACAGATCAGATTCTGTGTTGTG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAACCATCGAAATTCATCCGAGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysLeuPro 260
QY 839 ATTATGTTTGATGTCCTGATTCTCATGATCTGACCACTTCTCAATATTCTACCATTAAC 898
```

261	Db	IleWetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn	280
899	QY	TACAAGTGGAGCTTCGGGATATACTAGCGCTGTGTTGTTCCACCAATCATACTGTGAAT	958
281	Db	TyrluYtrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn	300
959	QY	CACACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCAGCTGTGAAAGCTCGACACA	1018
301	Db	HisThrTyrlValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro	320
1019	QY	GGACCTTGTCCGCCACGCCACACACACACAGACCTTCAAAACCCACCCCTCTTTAGGA	1078
321	Db	GlyProCysProProProProProProProProProProProProProProProSerLeuGly	340
1079	QY	CTGCTCGTGACAAACCCCTGGAGCTGAGTAGGATTCTCGATGAAAACGCCAGATTAAAC	1138
341	Db	ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn	360
1139	QY	AGATATGGCCATTTCAAGCCACCATCACAAATTGTAGAGGGAATCTTAGAGGTTAACTAC	1198
361	Db	ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle	380
1199	QY	ATCCAGATGACAGACGTCCTGATCCCGGTGCCATGGCCTAAAAGCTCCCTAATAGACTTT	1258
381	Db	IleGlnMetThrAspValLeuMetProValProTyrProGluSerSerLeuIleAspPhe	400
1259	QY	GTCGTGACCTGCCAAGGAGGACATTCACCGAGGCTGTGCACATCTTTCTGACCCCAACC	1318
401	Db	ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr	420
1319	QY	TGCAGATACCCAGAACACACAGCTGTCAGCCCTGTGATGCTGGATGGATGAGTGTCTGCTG	1378
421	Db	CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu	440
1379	QY	ACTGTGACAGCAACCTTCAATGGGTCTGGGACGTACTGTGTGACCTCCACCTGGGGGAT	1438
441	Db	ThrValArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp	460
1439	QY	GACACAAGCCTGGCTCTCACGAGACCCCTGATTTCTGTTCTGTGACAGACCCAGCCTCG	1498
461	Db	AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer	480
1499	QY	CGTTTAAGGATGGCAACAGTGCCTGTATCTCCGTTGGCTCTTGGCGCAATTTGTCACT	1558
481	Db	ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr	500
1559	QY	GTGATCTCCCTCTTCGTGTACAAAAAACACAGGAATACACCCAAATAGAAAAATAGTCTT	1618
501	Db	ValIleSerLeuLeuValTyrIlyIlyHisIlyGlyTyrAsnProIleGluAsnSerPro	520
1619	QY	GGGAATGTGGTCAGAGCAAGACCTGTGCTTCTTCAACCTGGCAAAAGCGGTGTC	1678
521	Db	GlyAsnValValArgSerIlyGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe	540
1679	QY	TTCCCGGGAAACCCAGGAAAAGATCCGGTACTTCAAAAAACCAAGATTTTAAAGGAGTTCT	1738
541	Db	PheProGlyIlyAsnGlnIlyIlyAsnProLeuIlyIlyAsnGlnGluPheIlyGlyValSer	560

## RESULT 13

```

US-10-219-535-42
; Sequence 42, Application US/10219535
; Publication NO. US20040044179A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.

```

```

1  APPLICANT: Wood, William I.
2  TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
3  TITLE OF INVENTION: ACIDS ENCODING THE SAME
4  FILE REFERENCE: P3530PIC60
5  CURRENT APPLICATION NUMBER: US/10/219,535
6  CURRENT FILING DATE: 2002-08-14
7  PRIOR APPLICATION NUMBER: 10/119,480
8  PRIOR FILING DATE: 2002-04-09
9  PRIOR APPLICATION NUMBER: 60/059113
10 PRIOR FILING DATE: 1997-09-17
11 PRIOR APPLICATION NUMBER: 60/062287
12 PRIOR FILING DATE: 1997-10-17
13 PRIOR APPLICATION NUMBER: 60/063549
14 PRIOR FILING DATE: 1997-10-28
15 PRIOR APPLICATION NUMBER: 60/064103
16 PRIOR FILING DATE: 1997-10-31
17 PRIOR APPLICATION NUMBER: 60/069873
18 PRIOR FILING DATE: 1997-12-17
19 PRIOR APPLICATION NUMBER: 60/078910
20 PRIOR FILING DATE: 1998-03-20
21 PRIOR APPLICATION NUMBER: 60/079294
22 PRIOR FILING DATE: 1998-03-25
23 PRIOR APPLICATION NUMBER: 60/079656
24 PRIOR FILING DATE: 1998-03-26
25 PRIOR APPLICATION NUMBER: 60/079728
26 PRIOR FILING DATE: 1998-03-27
27 Remaining Prior Application data removed - See File Wrapper or PALM.
28 NUMBER OF SEQ ID NOS: 246
29 SEQ ID NO 42.
30 LENGTH: 572
31 TYPE: PRT
32 ORGANISM: Homo Sapien
33 US-10-219-535-42

```

Alignment Scores:		
Pred. No.:	5.32e-268	572
Score:	2979.00	558
Percent Similarity:	97.55%	Conservative: 0
Best Local Similarity:	97.55%	Mismatches: 2
Query Match:	61.42%	Indels: 13
DP:	12	Gaps: 1

US 10-038-273-1 (1-2661) x US-10-219-535-42 (1-572)

QY	60	ATGGAATGCTCTACTACTATTTCTGGGATTTCTGCTCTGGCTGCAAGATTGCCACTTGAT	119
DB	1	MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuLeuLaalaArgLeuProLeuAsp	20
QY	120	GCCGCCAAACGATTTCATGTGTCTGGCAATGAAAGACCTTCTGTTACATGAGGAG	179
DB	21	AlaAlaIysArgPheHisAspValIleuGlyAsnGluArgProSerAlaTyrMetArgGlu	40
QY	180	CACAATCAATTAATGGCTGTCTTCGTGAAATGACTGGATGAAAACTCTACCCA	239
DB	41	HisAsnGlnLeuAsnGlyTyrSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro	60
QY	240	GTGTGGAAACGGGAGACATCAGCTGGMAAACTCCTGGAGGGAGCGCGTCAGCGC	299
DB	61	ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla	80
QY	300	GTCCTGACCACTGACTCACCAGCCCTCGTGGGCTCAAAATAACATTTGCGGTGAACCTG	359
DB	81	ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu	100
QY	360	ATATTCCTTCAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGACTGC	419
DB	101	IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys	120
QY	420	AGAAATGAGGCTGGTTTATCTGCTGATCCATGTTTACAACTGGACAGCATGFTCAGAG	479
DB	121	ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu	140
QY	480	GACAGTGAACGGGGAAAATGGCAACCGGCCAAAGCCATCAATACGCTCTTCCTGATCGGAAA	539

Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisHisAsnValPheProaspGlyLys 160  
QY 540 CTTTTCTCCACACCCCGATGGAGAAATGGAATTTTCATCTACCTCTCCACACATT 599  
Db 161 ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180  
QY 600 GGTCACTATTTCACAGAAATGGACGATGTTGAGTCAGAGAGTTCTGTGACACAGCCAT 659  
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200  
QY 660 GTGACACTTGGCCCTCAACTCATGGAAGTACTGTCTACAGAGACATGCAGCGGCATAT 719  
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220  
QY 720 GTTCCCATCGCACAGTGAAGATGTACGTGGTGAACAGATCAGATCTCTGTGTGG 779  
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240  
QY 780 ACTATGTTCCAGAGAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCC 838  
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260  
QY 839 ATTATGTTGATGCTCATCATCATCTAGCCACTTCTCTCAATATTCTACCAATTAAC 898  
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280  
QY 899 TACAGTGGAGCTTCGGGGAATACTAGCCCTGTTGTTTCCACCAATCATCTGTGAAT 958  
Db 281 TyrIleTyrSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300  
QY 959 CACAGTATGTGCTCAATGGAACCTTACGCTTACCTGCTGAAAGCTGCAGACCA 1018  
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320  
QY 1019 GGACTTGTCCGACCCGCCACCCACCCAGACCTTCAAAACCCACCCCTTCTTA--- 1075  
Db 321 GlyProCysProProProProProProProProProProSerLysProThrProSerLeuAla 340  
QY 1076 -----GGACTGCTGTGTGACAAACCCCTGTGAG 1102  
Db 341 ThrThrLeuLysSerTyrAspSerAsnThrProGlyProThrGlyAspAsnProLeuGlu 360  
QY 1103 CTGAGTAGGATTCCTGATGAAATGCCAGATTAACAGATATGCCACTTTCAGACCAAC 1162  
Db 361 LeuSerArgIleProAspGluAsnCysGlnIleAsnArgTyrGlyHisPheGlnAlaThr 380  
QY 1163 ATCAATTTGAGGGAACTTTAGAGTTAACAATCATCCAGATCAGACAGCTCTGTATG 1222  
Db 381 IleThrIleValGluGlyIleLeuGluValAsnIleIleGlnMetThrAspValLeuMet 400  
QY 1223 CCGTGTCCATGGCTGAAGCTCCCTAATAGACTTTGTGTGACCTGCCAAGGAGCAAT 1282  
Db 401 ProValProThrProGluSerSerLeuIleAspPheValValThrCysGlnGlySerIle 420  
QY 1283 CCCACGGAGTCTGTACCATCTATCTGACCCACCTGCGAGATCACCCAGACACAGTC 1342  
Db 421 ProThrGluValCysThrIleIleSerAspProThrCysGluIleThrGlnAsnThrVal 440  
QY 1343 TGAGCCCTGTGATGTGATGATGATGTGTGCTGACTGTGAGACGAACTTCAATGGG 1402  
Db 441 CysSerProValAspValAspGluMetCysLeuLeuThrValArgArgThrPheAsnGly 460  
QY 1403 TCTGGGACGTACTGTGTGAACCTCACCTGGGGATGACACAGCCTGTGCTCAGAGC 1462  
Db 461 SerGlyThrTyrCysValAsnLeuThrLeuGlyAspAspThrSerLeuAlaLeuThrSer 480  
QY 1463 ACCCTGATTTCTGTCTCTGACAGACCCAGCTCGCTTTAAGGATGGCAACAGTGGC 1522  
Db 481 ThrLeuIleSerValProAspArgAspProAlaSerProLeuArgMetAlaAsnSerAla 500  
QY 1523 CTGATCTCCGCTGGCTGGCCATATTGTCACTGTGATCTCCCTCTTGTGGTACAA 1582

Db 501 LeuIleSerValGlyCysLeuAlaIlePheValThrValIleSerLeuValTyrLys 520  
QY 1583 AAACACAAGAAATACAAACCAATAGAAATAGTCTCTGGGAATCTGTGACAGCAAGGC 1642  
Db 521 LysHisLysGluTyrAsnProIleGluAsnSerProGlyAsnValValArgSerLysGly 540  
QY 1643 CTGAGTGTCTTTCTCAACCGTCAAAAGCCGTGTTCTTCCCGGAAACACAGGAAAGGAT 1702  
Db 541 LeuSerValPheLeuAsnArgAlaLysAlaValPhePheProGlyAsnGlnGluLysAsp 560  
QY 1703 CGCTACTCAAAACCAAGATTTTAAAGAGTTTCT 1738  
Db 561 ProLeuLeuLysAsnGlnGluPheLysGlyValSer 572

## RESULT 14

US-10-232-230-42  
; Sequence 42, Application US/10232230  
; Publication No: US20040044180A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Gerritsen, Mary  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Philippe F.  
; APPLICANT: Watanabe, Colin L.  
; APPLICANT: Wood, William I.

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P3530P1C103  
CURRENT APPLICATION NUMBER: US/10/232,230  
CURRENT FILING DATE: 2002-08-29  
PRIOR APPLICATION NUMBER: 10/119,480  
PRIOR FILING DATE: 2002-04-09  
PRIOR APPLICATION NUMBER: 60/059113  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/062287  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/063549  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/064103  
PRIOR FILING DATE: 1997-10-31  
PRIOR APPLICATION NUMBER: 60/069873  
PRIOR FILING DATE: 1997-12-17  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 246  
; SEQ ID NO 42  
; LENGTH: 572  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-232-230-42

Alignment Scores:  
Pred. No.: 5,32e-268 Length: 572  
Score: 2979.00 Matches: 558  
Percent Similarity: 97.55% Conservative: 0  
Best Local Similarity: 97.55% Mismatches: 2  
Query Match: 61.42% Indels: 13  
DB: 12 Gaps: 1

US-10-039-272-1 (1-2661) x US-10-232-230-42 (1-572)



```
QY 60 ATGGAATGCTCTACTATTTCTCTGGGATTTCTGCTCGGTGCTGCAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCCGCCAAACGATTTCATGATGTCTGGGCAATGAAAGACCTTCTGCTTACATGAGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAATGCTGTCTTCATGATAAATGCTGAATGAATAAATCTTACCCA 239
Db 41 HisAsnGlnLeuAsnGlyTyrSerSerAspGluAsnAspTyrAsnGluLysLeuTyrPro 60
QY 240 GTGTGAAGCGGGAGACATGAGGTGGAAAACTCTCGAAGGGAGGCCGTGTGAGGCG 299
Db 61 ValTyrLysArgGlyAspMetArgTyrLysAsnSerTyrLysGlyArgValGlnAla 80
QY 300 GTCCTGACCACTGACTACCAAGCCCTCGTGGCTCAAAATATAACATTTGCGGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
QY 360 ATATTCCCTAGATGCCAAAGGAAGATGCCAATGCCAATAGTCTATCAGAAGAACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGlnLysAsnCys 120
QY 420 AGAAATGAGCTGGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGTGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTyrThrAlaTyrSerGlu 140
QY 480 GACAGTGAAGCGGAAATGGCACCGCCAAAGCCATATAAGCTCTTCCCTGATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
QY 540 CTTTTTCCTCACACCCCGGATGGAGAAGATTTTCATAGCTCTTCCACACACTT 599
Db 161 ProPheProHisHisProGlyTyrArgArgTyrPheIleTyrValPheHisThrLeu 180
QY 600 GTTCAGTATTTCCAGAAATGGACCATGTTTCAGTGAAGATTTCTGTGAACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGCCCTCAACTCATGGAAGTGAAGTCTGTACAGAAGACATGACGGGATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCAAGTGAAGATGTACGTGTAACAGATCAGATTTCTGTGTTTGG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAGAACGATCGAAATTCATCCGAGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTGATGCTCTGATTCATGATCCTAGCACCTTCTCAATTTATCTACCATTAAC 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTCGGGGATATCTGGCTGTTTGTTCACCAATCATCTGTGAAT 958
Db 281 TyrLysTyrSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGATATGCTCAATGAAACCTTCAGCTTAACTCACTGTGAAGCTGCAGCACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCCGACCGCACCAACCCAGACCTTCAAAACCCACCCCTCTTTA--- 1075
Db 321 GlyProCysProProProProProProProProProProProProProProProProPro 340
QY 1076 -----GGACCTGCTGTGTGACCAACCCCTCTGTGAG 1102
Db 341 ThrThrLeuLysSerTyrAspSerAsnThrProGlyProThrGlyAspAsnProLeuGlu 360
QY 1103 CTGAGTAGGATTCCTGATGAAACTGCCAGATTAAACAGATATGCCCACTTCAAGCCACC 1162
```

## RESULT 15

US-10-232-224-42

; Sequence 42, Application US/10232224

; Publication No. US20030065147A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Desnoyers, Luc

; APPLICANT: Gerritsen, Mary

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Smith, Victoria

; APPLICANT: Stephan, Jean-Philippe F.

; APPLICANT: Watanabe, Colin L.

; APPLICANT: Wood, William I.

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P350P1C111

; CURRENT APPLICATION NUMBER: US/10/232, 224

; CURRENT FILING DATE: 2002-08-29

; PRIOR APPLICATION NUMBER: 10/119,480

; PRIOR FILING DATE: 2002-04-09

; PRIOR APPLICATION NUMBER: 60/059113

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/062287

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/063549

; PRIOR FILING DATE: 1997-10-28

; PRIOR APPLICATION NUMBER: 60/064103



```

; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 42
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-232-224-42

Alignment Scores:
Pred. No.: 5,32e-268 Length: 572
Score: 2979.00 Matches: 558
Percent Similarity: 97.55% Conservative: 0
Best Local Similarity: 97.55% Mismatches: 2
Query Match: 61.42% Indels: 13
DB: 12 Gaps: 1

US-10-039-272-1 (1-2661) x US-10-232-224-42 (1-572)

QY 60 ATGGAAATGCTCTACTATTTCCTGGGATTTCCTCTGGCTGCAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCGGCCAAACGATTTCATCATGTCTGGCAATGAAGACCTTCCTTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheHisaspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAAATGGCTGCTCTCTGATGAAATGACTGGAATCAAACTCAACCA 239
Db 41 HisAsnGlnLeuAsnGlyTyrPheSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60
QY 240 GTGTGGGAAGCGGGAGACATGAGTGGGAAAACTCTGSAAGGGAGCGCTGTGACGGCG 299
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla 80
QY 300 GTCCTGACCACTGACTCACCAGCCCTCGTGGCTCAAAATATAACATTTGGCGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
QY 360 ATATTCCCTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAGAACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAATGAGGCTGTTTATCTGTCATCCATATGTTTACAACTGGACAGCATGTGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACAGTGACGGGAAAAATGGACCGGCCAAAGCCATCATACGCTTCCTCCGTATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
QY 540 CTTTTCCTCACCACCCCGGATGGAGAGATGGAAATTCATCTACGCTTCCACACACTT 599
Db 161 ProPheProHisHisProGlyTyrArgArgTrpAsnPhelIeTyrValPheHisThrLeu 180
QY 600 GGTCAATATTCAGAAATGGGACGAGTGTCACTGAGAGTTTCTGTGAACACAGCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTACACTTGGGCTCAACTCAGGAGTCACTGTCTACAGAAGACATGAGCGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220

```

```

QY 720 GTTCCCATCGCACAGTGAAGATGTGTACCTGGTAAACAGATCAGATTCTCTGTGTTGTG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAGAAGACGATCGAAATTCATCCGACGAAACCTTCCG-AAAATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTTTCATGCTCCTGATTTCATGATCCCTAGCCACTTCCTCAATATTTCTACATTAA 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGGATATACTGGCCTCTTTGTTTCCCAACCAATATCTGTGAAT 958
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTCTCAATGAACCTTCAGCCTTAACTCACTGTGAAGCTGCAGACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCCGCCACCGCCACCCACCCAGACCTTCAAAACCCACCCCTTCTTTA--- 1075
Db 321 GlyProCysProProProProProProProProProProProProProProSerLeuAla 340
QY 1076 -----GGACCTGCTGCTGACAAACCCCTGGAG 1102
Db 341 ThrThrLeuLysSerTyrAspSerAsnThrProGlyProThrGlyAspAsnProLeuGlu 360
QY 1103 CTGAGTAGGATTCCTGATGAAACCTGCCAGATTAAACAGATATGGCCACTTCAAGCCACC 1162
Db 361 LeuSerArgIleProAspGluAsnCysGlnIleAsnArgTyrGlyHisPheGlnAlaThr 380
QY 1163 ATCACAATGTAGAGGGAATCTTAGAGTTAAATCATCATCCAGATCAGACAGCTCTCTGATG 1222
Db 381 IleThrIleValGluGlyIleLeuGluValAsnIleIleGlnMetThrAspValLeuMet 400
QY 1223 CCGTGCATCGCTGAAAGCTCCCTAATAGACTTTTGTGTCACCTGCCAGCCAGGGAGCATT 1282
Db 401 ProValProTyrProGluSerSerLeuIleAspPheValValThrCysGlnGlySerIle 420
QY 1283 CCCAGGAGTCTGTACCATCATTTCTGACCCCACTCCGAGATCACCAGAAACACAGTC 1342
Db 421 ProThrGluValCysThrIleSerAspProThrCysGluIleThrGlnAsnThrVal 440
QY 1343 TCCAGCCCTGGGATGTGATGAGATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1402
Db 441 CysSerProValAspValAspGluMetCysLeuLeuThrValArgArgThrPheAsnGly 460
QY 1403 TCTGGGACGTACTGTGTGAACCTCACCTCGGGGATGACACAAGCCTGGCTCTCACGAGC 1462
Db 461 SerGlyThrTyrCysValAsnLeuThrLeuGlyAspAspThrSerLeuAlaLeuThrSer 480
QY 1463 ACCCTGATTTCTGTTCTCTGACAGACCCAGCTCGCTGCTTAAAGATGGCAAAACAGTGCC 1522
Db 481 ThrLeuIleSerValProAspArgAspProAlaSerProLeuArgMetAlaAsnSerAla 500
QY 1523 CTGATCTCCGCTGGCTCTGGCCATATTTGCTCACTGTGATCTCCCTCTGCTGCTGCTGCTG 1582
Db 501 LeuIleSerValGlyCysLeuAlaIlePheValThrValIleSerLeuLeuValTyrLys 520
QY 1583 AAACACAGGAATACAAACCAATAGAAAAATAGTCTGGGAATGTGGTCAGAGCAAGAGGC 1642
Db 521 LysHisLysGluTyrAsnProIleGluAsnSerProGlyAsnValValArgSerLysGly 540
QY 1643 CTGAGTGTCTTCTCAACCGTGCAAAAGCGGTGTTCTTCCCGGAAACCAAGGAAAGGAT 1702
Db 541 LeuSerValPheLeuAsnArgAlaLysAlaValPhePheProGlyAsnGlnGluLysAsp 560
QY 1703 CCCTACTCAAAAACCAAGAAATTTAAGGAGTTTCT 1738
Db 561 ProLeuLysAsnGlnGluPheLysGlyValSer 572

```

us-10-039-272-1.rapb

Mon Sep 13 10:28:10 2004

Search completed: September 10, 2004, 16:38:13  
Job time : 613.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2004, 07:08:49 ; Search time 3913 Seconds  
(without alignments)  
3421.641 Million cell updates/sec

```

Title: US-10-039-272-1
Perfect score: 2661
Sequence: 1 cggcagcagggcccagagga.....aaaaaaaaaaaaaaaaaaaaa 2661

```

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3304383 seqs, 2515761380 residues  
Total number of hits satisfying chosen parameters

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 s

```

Database : Published Applications NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARY

Result No.	Query %			Description	
	Score	Match	Length	ID	Description
1	2661	100.0	2661	13	US-10-463-106-1
2	2661	100.0	2661	14	US-10-039-272-1
3	2559.2	96.2	2352	15	US-10-198-846-10843
4	2547.8	95.7	2669	10	US-09-525-9788-82
5	2547.8	95.7	2669	13	US-10-342-687-779
6	2547.8	95.7	2669	13	US-10-172-118-779
7	2547.8	95.7	2669	16	US-10-295-027-1041
8	2547.8	95.7	2669	17	US-10-450-826-92
9	2531.6	95.1	2666	15	US-10-084-817-169
10	2485	93.4	2683	13	US-10-219-535-41
11	2485	93.4	2683	13	US-10-332-230-41
12	2485	93.4	2683	13	US-10-432-224-41
13	2485	93.4	2683	15	US-10-227-684-41
14	2485	93.4	2683	15	US-10-230-163-41

15	2485	93.4	2683	15	US-10-230-338-41	Sequence 41, Appl
16	2485	93.4	2683	15	US-10-218-631-41	Sequence 41, Appl
17	2485	93.4	2683	15	US-10-230-414-41	Sequence 41, Appl
18	2485	93.4	2683	15	US-10-216-159A-41	Sequence 41, Appl
19	2485	93.4	2683	15	US-10-218-849-41	Sequence 41, Appl
20	2485	93.4	2683	15	US-10-227-873-41	Sequence 41, Appl
21	2485	93.4	2683	15	US-10-227-883-41	Sequence 41, Appl
22	2485	93.4	2683	15	US-10-219-076-41	Sequence 41, Appl
23	2485	93.4	2683	15	US-10-230-434-41	Sequence 41, Appl
24	2485	93.4	2683	15	US-10-219-003-41	Sequence 41, Appl
25	2485	93.4	2683	15	US-10-219-075-41	Sequence 41, Appl
26	2485	93.4	2683	15	US-10-219-464-41	Sequence 41, Appl
27	2485	93.4	2683	15	US-10-219-466-41	Sequence 41, Appl
28	2485	93.4	2683	15	US-10-219-479-41	Sequence 41, Appl
29	2485	93.4	2683	15	US-10-219-481-41	Sequence 41, Appl
30	2485	93.4	2683	15	US-10-230-260-41	Sequence 41, Appl
31	2485	93.4	2683	15	US-10-232-231-41	Sequence 41, Appl
32	2485	93.4	2683	15	US-10-232-233-41	Sequence 41, Appl
33	2485	93.4	2683	15	US-10-216-165-41	Sequence 41, Appl
34	2485	93.4	2683	15	US-10-218-956-41	Sequence 41, Appl
35	2485	93.4	2683	15	US-10-219-468-41	Sequence 41, Appl
36	2485	93.4	2683	15	US-10-219-478-41	Sequence 41, Appl
37	2485	93.4	2683	15	US-10-219-536-41	Sequence 41, Appl
38	2485	93.4	2683	15	US-10-233-205-41	Sequence 41, Appl
39	2485	93.4	2683	15	US-10-219-072-41	Sequence 41, Appl
40	2485	93.4	2683	15	US-10-219-470-41	Sequence 41, Appl
41	2485	93.4	2683	15	US-10-219-474-41	Sequence 41, Appl
42	2485	93.4	2683	15	US-10-219-524-41	Sequence 41, Appl
43	2485	93.4	2683	15	US-10-219-528-41	Sequence 41, Appl
44	2485	93.4	2683	15	US-10-227-880-41	Sequence 41, Appl
45	2485	93.4	2683	15	US-10-227-881-41	Sequence 41, Appl

## ALIGNMENTS

```

RESULT 1
US-10-463-106-1
; Sequence 1, Application US/10463106
; Publication No. US20030202938A1
; GENERAL INFORMATION:
; APPLICANT: RAMESHWAR, Pranela
; TITLE OF INVENTION: HEMATOPOIETIC GROWTH FACTOR INDUCIBLE NEUROKININ-1 GENE
; FILE REFERENCE: 267/033 University of Medicine & Dentistry of New Jersey
; CURRENT APPLICATION NUMBER: US/10/463.106
; CURRENT FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: US 10/039,272
; PRIOR FILING DATE: 2001-10-20
; PRIOR APPLICATION NUMBER: US 60/241,881
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 2661
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-463-106-1

```

	Query Match	100.0%;	Score 2661;	DB 13;	Length 2661;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 2661;	Conservative	0;	Mismatches	0;
				Indels	Gaps
				0;	0;
QY	1	CGGCACGAGGGCCAGAGGAATAAGTTAACTTTGGTGCCTGCGTCCGTGAGATTTCAGCA	60		
Db	1	CGGCACGAGGGCCAGAGGAATAAGTTAACTTTGGTGCCTGCGTCCGTGAGATTTCAGCA	60		
QY	61	TGGAATGTCTCTACTATTTCCTGGGATTTCTGCTCGCTGCAAGATTCGCCACTTGATG	120		
Db	61	TGGAATGTCTCTACTATTTCCTGGGATTTCTGCTCGCTGCAAGATTCGCCACTTGATG	120		
QY	121	CGGCCAAACGATTTTCATGATGTGCTGGCAATGAAAGACCTTCTGCTTACATGAGGGAGC	180		
Db	121	CGGCCAAACGATTTTCATGATGTGCTGGCAATGAAAGACCTTCTGCTTACATGAGGGAGC	180		





1441 CACAAGCCTGGCTCTCAGGACACCCCTGATTTCTGTTCTTCTGACAGAGACCCAGCCTCGCC 1500  
1501 TTTTAAGGATGGCAACAGTCCCTGATCTCCGTTGGTGCTGTTGGCCATATTTGTCACTGT 1560  
1501 TTTTAAGGATGGCAACAGTCCCTGATCTCCGTTGGTGCTGTTGGCCATATTTGTCACTGT 1560  
1561 GATCTCCCTCTTGGTGACAAAACACAAAGGAATACAAACCAATAGAAAAATAGTCCCTGG 1620  
1561 GATCTCCCTCTTGGTGACAAAACACAAAGGAATACAAACCAATAGAAAAATAGTCCCTGG 1620  
1621 GAATGTGTGACAAAGGCTGAGTGTCTTTCTCAACCGTGCAAAAGCGGTGTTCTT 1680  
1621 GAATGTGTGACAAAGGCTGAGTGTCTTTCTCAACCGTGCAAAAGCGGTGTTCTT 1680  
1681 CCCGGAAACACAGGAAAGGATCCGCTACTCAAAACCAAGAAATTTAAAGGAGTTCTTA 1740  
1681 CCCGGAAACACAGGAAAGGATCCGCTACTCAAAACCAAGAAATTTAAAGGAGTTCTTA 1740  
1741 AATTTGACCTTGTGTTGAGCTCACTTTTCAGTGCCATTCATGATGTGCTGGAG 1800  
1741 AATTTGACCTTGTGTTGAGCTCACTTTTCAGTGCCATTCATGATGTGCTGGAG 1800  
1801 TGGCTATTAACTTTTTTCTTAAAGATTTATTTTAAATAGATATTTGTTGGTGGGAG 1860  
1801 TGGCTATTAACTTTTTTCTTAAAGATTTATTTTAAATAGATATTTGTTGGTGGGAG 1860  
1861 TTGAATTTTTTATAGTTAAATGTCATTTTATAGATGGGAGAGGATTTATCTGAGGC 1920  
1861 TTGAATTTTTTATAGTTAAATGTCATTTTATAGATGGGAGAGGATTTATCTGAGGC 1920  
1921 AGCTTCAAGCTTGTGAACTGATTAAGCAACTTAGCAAGCTTCTTTTCAATATTT 1980  
1921 AGCTTCAAGCTTGTGAACTGATTAAGCAACTTAGCAAGCTTCTTTTCAATATTT 1980  
1981 TTTATGTTTCACTTATAAGCTCTTAGTAACTAGTAGGATAGAAACACTGTGTCGAG 2040  
1981 TTTATGTTTCACTTATAAGCTCTTAGTAACTAGTAGGATAGAAACACTGTGTCGAG 2040  
2041 GTAAGGAGAGAAGCTACTATTGATAGAGCTTAACCGAGTTAACTGCAAGAGAGCGG 2100  
2041 GTAAGGAGAGAAGCTACTATTGATAGAGCTTAACCGAGTTAACTGCAAGAGAGCGG 2100  
2101 GATACCTTCACTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160  
2101 GATACCTTCACTTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160  
2161 TCATGTTCCAGCTTAAGTAACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220  
2161 TCATGTTCCAGCTTAAGTAACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220  
2221 ATAAAGGCGGCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280  
2221 ATAAAGGCGGCAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280  
2281 TCTCATAAATGGGTTGGGATTTTGGTGACAACTTCTGTTGCTGCTGCTGCTGCTGCTGCTG 2340  
2281 TCTCATAAATGGGTTGGGATTTTGGTGACAACTTCTGTTGCTGCTGCTGCTGCTGCTGCTG 2340  
2341 TGATATTCAATATTTCAATTTATTCATGGAATTTAGTATGCTGCTGCTGCTGCTGCTGCTG 2400  
2341 TGATATTCAATATTTCAATTTATTCATGGAATTTAGTATGCTGCTGCTGCTGCTGCTGCTG 2400  
2401 ATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2460  
2401 ATGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2460  
2461 TTGAAATCAAAATATTAAGCTTTTCCAAAATTTGCTCCCTGCTGCTGCTGCTGCTGCTGCTG 2520  
2461 TTGAAATCAAAATATTAAGCTTTTCCAAAATTTGCTCCCTGCTGCTGCTGCTGCTGCTGCTG 2520  
2521 ATCAGTAGGATTTCCCTCTGTTTGGAACTTAAACCAATTTATATATGTTAGACAAGAC 2580  
2521 ATCAGTAGGATTTCCCTCTGTTTGGAACTTAAACCAATTTATATATGTTAGACAAGAC 2580

2581 ATTTTTTTTTTCTTCTGAAATGAGGAGAGACACAAATGAGGAGAGAGAGAGAGAGAGAGAG 2640  
2581 ATTTTTTTTTTCTTCTGAAATGAGGAGAGAGACACAAATGAGGAGAGAGAGAGAGAGAGAGAG 2640  
2641 AAAAAAAAAAAAAAAAAAAAAA 2661  
2641 AAAAAAAAAAAAAAAAAAAAAA 2661

RESULT 3  
US-10-198-846-10843  
; Sequence 10843, Application US/10198846  
; Publication No. US2003009974A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Xu, Yongyao  
; APPLICANT: Wang, Youzhen  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF BREAST CANCER  
; FILE REFERENCE: MRI-049  
; CURRENT APPLICATION NUMBER: US/10/198,846  
; CURRENT FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/306,220  
; PRIOR FILING DATE: 2001-07-18  
; NUMBER OF SEQ ID NOS: 14084  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10843  
; LENGTH: 2952  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 2950..2951, 2952  
; OTHER INFORMATION: n = A,T,C or G  
US-10-198-846-10843

Query Match 96.2%; Score 2559.2; DB 15; Length 2952;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 2606; Conservative 0; Mismatches 23; Indels 3; Gaps 3;

28 AACCTTGGTCCCTGCTGCTGAGATTCAGCATGGAATGCTCTACTATTTCCTGGAT 87  
187 ACCTTGAGTCTGCTGCTGCTGAGATTCAGCATGGAATGCTCTACTATTTCCTGGAT 246  
88 TTCTGCTCTGCTGCTGCTGAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 147  
247 TTCTGCTCTGCTGCTGCTGAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 306  
148 GCAATGAAAGACCTTCTGCTTACATGAGGAGACCAATCAATTAATGCTGCTGCTGCTG 207  
307 GCAATGAAAGACCTTCTGCTTACATGAGGAGACCAATCAATTAATGCTGCTGCTGCTG 366  
208 ATGAAATGATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 267  
367 ATGAAATGATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 426  
268 AAAATCTCTGAAAGGAGGCGCTGTCAGCGGCTGTCAGCGGCTGTCAGCGGCTGTCAG 327  
427 AAAATCTCTGAAAGGAGGCGCTGTCAGCGGCTGTCAGCGGCTGTCAGCGGCTGTCAG 486  
328 TGGGCTCAATATATCAATTTTGGGCTGAACTGATATTTCCCTAGATGCTGCTGCTGCTG 387  
487 TGGGCTCAATATATCAATTTTGGGCTGAACTGATATTTCCCTAGATGCTGCTGCTGCTG 546  
388 CCAATGGAACATAGTCTATGAGAAAGAACTGCAAGAAAGAGCTGCTGCTGCTGCTGCTG 447  
547 CCAATGGAACATAGTCTATGAGAAAGAACTGCAAGAAAGAGCTGCTGCTGCTGCTGCTG 606  
448 CATATGTTTCAACTGGACAGCATGCTGAGAGGACAGTGGAGGGGAAATGACCGGCGC 507

Db 607 CGTATGTTTACAACTGGACAGCATGGTCAGAGACAGTGCACGGGGAAAAATGGCACCGGCC 566  
QY 508 AAAGCCATCATAAAGTCCTCCCTGATGGGAACCTTTTCTCACCACCCCGGATGGAGAA 567  
Db 667 AAAGCCATCATAAAGTCCTCCCTGATGGGAACCTTTTCTCACCACCCCGGATGGAGAA 726  
QY 568 GATGGAATTTCACTACAGTCCTCCACACACACTTGGTCAGTATTTCCAGAAAATTTGGGACGAT 627  
Db 727 GATGGAATTTCACTACAGTCCTCCACACACTTGGTCAGTATTTCCAGAAAATTTGGGACGAT 786  
QY 628 GTTCAGTGAGAGTTTCTGTGAAACACAGCCAAATGTGACACTTGGGCTCAACTCATGGAAG 687  
Db 787 GTTCAGTGAGAGTTTCTGTGAAACACAGCCAAATGTGACACTTGGGCTCAACTCATGGAAG 846  
QY 688 TGACGTGCTACAGAGACATGACGGGCATATGTTCCCATCCACAAAGTGAAGATGTGT 747  
Db 847 TGACGTGCTACAGAGACATGACGGGCATATGTTCCCATCCACAAAGTGAAGATGTGT 906  
QY 748 AGCTGGTAAACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAAAGAACGATCGAAAT 807  
Db 907 AGCTGGTAAACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAAAGAACGATCGAAAT 966  
QY 808 CATCCGACGAACCTTCC-CAAGATCTCCCAATATGTTGATGTCTGTGATTCATGATC 866  
Db 967 CATCCGACGAACCTTCCCAAGATCTCCCAATATGTTGATGTCTGTGATTCATGATC 1026  
QY 867 CTAGCCACTCTCTCAATTAATCTACCAATTAATCTACAGTGGAGCTTCGGGGTAATACTG 926  
Db 1027 CTAGCCACTCTCTCAATTAATCTACCAATTAATCTACAGTGGAGCTTCGGGGTAATACTG 1086  
QY 927 GCCTGTTTGTTCACCAATCATACTGTGAATCACACGATATGCTCAATGAACTTCA 986  
Db 1087 GCCTGTTTGTTCACCAATCATACTGTGAATCACACGATATGCTCAATGAACTTCA 1146  
QY 987 GCCTTAACCTCACTGTGAAGCTGAGACACAGGACCTTGTCCGCCACCGCCACCAACC 1046  
Db 1147 GCCTTAACCTCACTGTGAAGCTGAGACACAGGACCTTGTCCGCCACCGCCACCAACC 1206  
QY 1047 CCAGACCTTCAAAACCCACCTTCTTTAGGACCTGCTGGTGACAAACCCCTGGAGCTGA 1106  
Db 1207 CCAGACCTTCAAAACCCACCTTCTTTAGGACCTGCTGGTGACAAACCCCTGGAGCTGA 1266  
QY 1107 GTAGGATTCCTGATGAAACTGCCAGATTAACAGATATGGCCACTTTTCAAGCCACCATCA 1166  
Db 1267 GTAGGATTCCTGATGAAACTGCCAGATTAACAGATATGGCCACTTTTCAAGCCACCATCA 1326  
QY 1167 CAATGTGAGAGGAACTTTGAGGTAAATCCAGATGACAGACCTCTCTGATGCCGG 1226  
Db 1327 CAATGTGAGAGGAACTTTGAGGTAAATCCATCATCCAGATGACAGACCTCTCTGATGCCGG 1386  
QY 1227 TGCCATGGCCCTGAAAGCTCCCTAATAGACTTTGTGTCGACCTGCCAAGGGAGCATTCCTCA 1286  
Db 1387 TGCCATGGCCCTGAAAGCTCCCTAATAGACTTTGTGTCGACCTGCCAAGGGAGCATTCCTCA 1446  
QY 1287 CGGAGGCTGTACCATCATTTTCTGACCCCACTGCGAGATCACCCAGAACACAGTCTGCA 1346  
Db 1447 CGGAGGCTGTACCATCATTTCTGACCCCACTGCGAGATCACCCAGAACACAGTCTGCA 1506  
QY 1347 GCCCTGTGGATGTGATGAGATGTCTGTGACTGTGAGACGAACCTTCAATGGGCTG 1406  
Db 1507 GCCCTGTGGATGTGATGAGATGTCTGTGACTGTGAGACGAACCTTCAATGGGCTG 1566  
QY 1407 GGAGTACTGTGTGAACTGACCTGCGGGGATGACAAAGCTGCTCTCAAGAGCACCC 1466  
Db 1567 GGAGTACTGTGTGAACTGACCTGCGGGGATGACAAAGCTGCTCTCAAGAGCACCC 1626  
QY 1467 TGAATTTCTGTTCTGACAGACACCCAGCTCGCCCTTTAAGGATGGCAACAGTGCCTGA 1526  
Db 1627 TGAATTTCTGTTCTGACAGACACCCAGCTCGCCCTTTAAGGATGGCAACAGTGCCTGA 1686  
QY 1527 TCTCCGTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAAAAC 1586  
Db 1687 TCTCCGTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAAAAC 1746

QY 1587 ACAAGGAATACAAACCAATAGAAAAATAGTCTCTGGGAATCTGTCAGAAGCAAAAGGCTTGA 1646  
Db 1747 ACAAGGAATACAAACCAATAGAAAAATAGTCTCTGGGAATCTGTCAGAAGCAAAAGGCTTGA 1806  
QY 1647 GGTGCTTTCTCAACCGTGCAAAAGCGGTCTTCTCCCGGAAACCAAGGAAAGGATCCCG 1706  
Db 1807 GGTGCTTTCTCAACCGTGCAAAAGCGGTCTTCTCCCGGAAACCAAGGAAAGGATCCCG 1866  
QY 1707 TACTCAAAAACCAAGAAATTTAAAGGAGTTTCTTAAATTTCCGACCTTGTCTTCTGAAGCTCA 1766  
Db 1867 TACTCAAAAACCAAGAAATTTAAAGGAGTTTCTTAAATTTCCGACCTTGTCTTCTGAAGCTCA 1926  
QY 1767 CTTTTCAGTGCCCAATGATGTGATGTGCTGAGTGGCTTAATAACCTTTTTTCTCTAAAG 1826  
Db 1927 CTTTTCAGTGCCCAATGATGTGATGTGCTGAGTGGCTTAATAACCTTTTTTCTCTAAAG 1986  
QY 1827 ATTAATGTTAAATAGATATGTTGGTTTGGGAAAGTTGAATTTTTTATAGTTTAAATGTCA 1886  
Db 1987 ATTAATGTTAAATAGATATGTTGGTTTGGGAAAGTTGAATTTTTTATAGTTTAAATGTCA 2046  
QY 1887 TTTTAGAGATGGGAGAGGGATTAATCTGACGACGCTTTCAGCCATGTTGTGAAAATGAT 1946  
Db 2047 TTTTAGAGATGGGAGAGGGATTAATCTGACGACGCTTTCAGCCATGTTGTGAAAATGAT 2106  
QY 1947 AAAACCACTTAGCAAGGCTTCTTTTCAATTAATTTTTATGTTTCACTTATAAAGTCTTAG 2006  
Db 2107 AAAACCACTTAGCAAGGCTTCTTTTCAATTAATTTTTTATGTTTCACTTATAAAGTCTTAG 2166  
QY 2007 GTAACTACTAGGATAGAAACACTGTGTCGAGAGTAAGGAGAGAGCTACTATTGATTA 2066  
Db 2167 GTAACTAGTAAAGTAAACACTGTGTCGAGAGTAAGGAGAGAGCTACTATTGATTA 2226  
QY 2067 GAGCCTAACCCAGGTTAACTGCAAGAGAGCGGGATCTTTCAGCTTTCATGTAAGCTG 2126  
Db 2227 GAGCCTAACCCAGGTTAACTGCAAGAGAGCGGGATCTTTCAGCTTTCATGTAAGCTG 2286  
QY 2127 TATGCAATAAGCCCAATGATGATGCTTCTAAGATCATGTTTCCAAGCTAACTGAATCCCA 2186  
Db 2287 TATGCAATAAGCCCAATGATGATGCTTCTAAGATCATGTTTCCAAGCTAACTGAATCCCA 2346  
QY 2187 CTTCAATACACACTCATGAACCTCTGTGATGGAACAAATAACAGAGCCCAAGCTGTGGTATGA 2246  
Db 2347 CTTCAATACACACTCATGAACCTCTGTGATGGAACAAATAACAGAGCCCAAGCTGTGGTATGA 2406  
QY 2247 TGTGCACACTTGTAGACTCAGAAAAATACTACTCTCATAAATGGGTGGGAGTATTTTG 2306  
Db 2407 TGTGCACACTTGTAGACTCAGAAAAATACTACTCTCATAAATGGGTGGGAGTATTTTG 2466  
QY 2307 GTGCAACCTTACTTGTGCTGAGTGAAGGAATGATTAATCATATATTCATTTATTCCA 2366  
Db 2467 GTGCAACCTTACTTGTGCTGAGTGAAGGAATGATTAATCATATATTCATTTATTCCA 2526  
QY 2367 TGGACATTTAGTGTAGTCTTTTATATACAGGACATGATGCTGAGTGACACTCTGTGTGA 2426  
Db 2527 TGGACATTTAGTGTAGTCTTTTATATACAGGACATGATGCTGAGTGACACTCTGTGTGA 2586  
QY 2427 TATTTCCAAATTTTGTATAGTCTGCACATATTTGAAATCAAAATATTAAGACTTTTC 2486  
Db 2587 TATTTCCAAATTTTGTATAGTCTGCACATATTTGAAATC-ATATATTAAAGACTTTTC 2645  
QY 2487 AAAAATTTGTCTCTGTTTTCATGGCAACTTGCATGAGTAAGGATTTCCCTCTGTTTG 2546  
Db 2646 AAAGATGAGTCCCTGTTTTCATGGCAACTTGCATGAGTAAGGATTTCACTCTGTTTG 2705  
QY 2547 GAACTTAAACCAATTTACTATATGTTAGACAGCAATTTTTTTTTTCTCTCTCTGAAAAA 2606  
Db 2706 TAACTAAACCAATTTACTATATGTTAGACATGCAATTTTTTTCTCTCTCTCTGAAAAA 2765  
QY 2607 -AAAAATGAGGAGAGACAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 2657  
Db 2766 TAAAGTCTGGAGAGACAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 2817

## RESULT 4

US-09-525-978B-82

; Sequence 82, Application US/09525978B

; Publication No. US20030049722A1

; GENERAL INFORMATION:

; APPLICANT: Murray, Richard

; APPLICANT: Caras, Ingrid W.

; APPLICANT: Hevezi, Peter

; APPLICANT: Wilson, Keith

; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSING MACROPHAGE DEVELOPMENT

; TITLE OF INVENTION: RELATED DISORDERS, COMPOSITIONS, AND METHODS OF

; TITLE OF INVENTION: SCREENING FOR MACROPHAGE DEVELOPMENT MODULATORS

; FILE REFERENCE: A-67413-1/DJB/JJD

; CURRENT APPLICATION NUMBER: US/09/525, 978B

; CURRENT FILING DATE: 2000-03-15

; PRIOR APPLICATION NUMBER: USSN 60/124,530

; PRIOR FILING DATE: 1999-03-15

; NUMBER OF SEQ ID NOS: 83

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 82

; LENGTH: 2669

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-525-978B-82

Query Match 95.7%; Score 2547.8; DB 10; Length 2669;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 2591; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

QY	28	ACCTTGGTGCCTGCGTGGAGAAATCAGCATGAATGTCTACTATTTCTCGGGAT	87
DB	60	ACCTGAGTGCCTGCGTGGAGAAATCAGCATGAATGTCTACTATTTCTCGGGAT	119
QY	88	TTCTGCTCCTGGTGCAGAAATGCCACTTGAATCCGCCAAACGATTTATGATGCTGG	147
DB	120	TTCTGCTCCTGGTGCAGAAATGCCACTTGAATCCGCCAAACGATTTATGATGCTGG	179
QY	148	GCAATGAAGACCTTCTGCTTACATGAGGAGACAAATCAATTTAAATGGTGTCTTCTG	207
DB	180	GCAATGAAGACCTTCTGCTTACATGAGGAGACAAATCAATTTAAATGGTGTCTTCTG	239
QY	208	ATGAAATGACTGGAATGAAATCTTACCAGTGTGGAAGCGGGAGACATGAGGTGGA	267
DB	240	ATGAAATGACTGGAATGAAATCTTACCAGTGTGGAAGCGGGAGACATGAGGTGGA	299
QY	268	AAAACTCCTGGAAGGAGCGGTGTGTCAGCGGTCTCTGACAGTCAACAGCCCTCG	327
DB	300	AAAACTCCTGGAAGGAGCGGTGTGTCAGCGGTCTCTGACAGTCAACAGCCCTCG	359
QY	328	TGGGCTCAATATTAACATTTGCGGTGAACCTGATATTCCTAGATGCCAAAGGAGATG	387
DB	360	TGGGCTCAATATTAACATTTGCGGTGAACCTGATATTCCTAGATGCCAAAGGAGATG	419
QY	388	CCATGCAACATAGTCTATGAGAAAGTGCAGAAATGAGGTGTTTATCTGCTGATC	447
DB	420	CCATGCAACATAGTCTATGAGAAAGTGCAGAAATGAGGTGTTTATCTGCTGATC	479
QY	448	CATATGTTTAACTGACAGCATGGTTCAGAGGACAGTGCAGGGAATAATGGCACCGCC	507
DB	480	CATATGTTTAACTGACAGCATGGTTCAGAGGACAGTGCAGGGAATAATGGCACCGCC	539
QY	508	AAAGCCATCAATAGCTTCTCCCTGATGGAAACCTTTTCCCTACCAACCCCGGATGAGAA	567
DB	540	AAAGCCATCAATAGCTTCTCCCTGATGGAAACCTTTTCCCTACCAACCCCGGATGAGAA	599
QY	568	GATGGAATTTTCACTGCTTCTCCACACATTTGCTGAGTATTTCCAGAAATTTGGGAGAT	627
DB	600	GATGGAATTTTCACTGCTTCTCCACACATTTGCTGAGTATTTCCAGAAATTTGGGAGAT	659
QY	628	GTTTCAGTGAGAGTTTCTGTGAACACAGCCAAATGTGACATTTGGGCTCACTATGGAAG	687
DB	660	GTTTCAGTGAGAGTTTCTGTGAACACAGCCAAATGTGACATTTGGGCTCACTATGGAAG	719

QY	688	TGACTGTCTACAGAGACATGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT	747
DB	720	TGACTGTCTACAGAGACATGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT	779
QY	748	ACGTGTAACAGATCAGATTTCTCTGTGTTTGTGATATGTTTCCAGAAAGACGATCGAAAT	807
DB	780	ACGTGTAACAGATCAGATTTCTCTGTGTTTGTGATATGTTTCCAGAAAGACGATCGAAAT	839
QY	808	CATCCGAGAAACCTTCC-CAAAGATCTCCCATATGTTTGTGATGTCCTGATTCATGATC	866
DB	840	CATCCGAGAAACCTTCTCAAAGATCTCCCATATGTTTGTGATGTCCTGATTCATGATC	899
QY	867	CTAGCCACTTCTCAATATTTCTACCAATTAACATAAGTGGAGCTTCGGGGATAATACTG	926
DB	900	CTAGCCACTTCTCAATATTTCTACCAATTAACATAAGTGGAGCTTCGGGGATAATACTG	959
QY	927	GCCTGTTGTTTCCCAACCAATCATATCTGTAATTCACAGTATGTCATGTAAGACCTTCA	986
DB	960	GCCTGTTGTTTCCCAACCAATCATATCTGTAATTCACAGTATGTCATGTAAGACCTTCA	1019
QY	987	GCCTTAACTCACTGTGAAGCTGACGACACAGGACCTTGTCCGCCACCGCCACCAAC	1046
DB	1020	GCCTTAACTCACTGTGAAGCTGACGACACAGGACCTTGTCCGCCACCGCCACCAAC	1079
QY	1047	CCAGACCTTCAAAACCCACCCCTTCTTTAGGACCTGCTGCTGTGACAAACCCCTTGGAGCTGA	1106
DB	1080	CCAGACCTTCAAAACCCACCCCTTCTTTAGGACCTGCTGCTGTGACAAACCCCTTGGAGCTGA	1139
QY	1107	GTAGGATTCCTGATGAAAACTGCGCAATTAACAGATATGCCACTTTCAAGCCACCATCA	1166
DB	1140	GTAGGATTCCTGATGAAAACTGCGCAATTAACAGATATGCCACTTTCAAGCCACCATCA	1199
QY	1167	CAATTTGAGGGAACTTTAGAGGTTAAACATCATCCAGATGACAGACGCTCTGATGCCGG	1226
DB	1200	CAATTTGAGGGAACTTTAGAGGTTAAACATCATCCAGATGACAGACGCTCTGATGCCGG	1259
QY	1227	TGCCATGGCTGAAAGCTCCCTAAATAGACTTTGCTGTGACCTGCCAAGGGAGCATTTCCCA	1286
DB	1260	TGCCATGGCTGAAAGCTCCCTAAATAGACTTTGCTGTGACCTGCCAAGGGAGCATTTCCCA	1319
QY	1287	CGAGGCTGTACCATCAATTTCTGACCCCACTGCGAGATCACCCAGAACACAGTCTGCA	1346
DB	1320	CGAGGCTGTACCATCAATTTCTGACCCCACTGCGAGATCACCCAGAACACAGTCTGCA	1379
QY	1347	GCCTCTGGATGTGATGAGTGTCTGCTGATCTGTGAGACCAACCTTCAATGGGTCTG	1406
DB	1380	GCCTCTGGATGTGATGAGTGTCTGCTGATCTGTGAGACCAACCTTCAATGGGTCTG	1439
QY	1407	GGACGTACTGTGTGAACCTTCAACCCCTGGGGATGACAAAGCTTGGCTTCTCAGAGCACCC	1466
DB	1440	GGACGTACTGTGTGAACCTTCAACCCCTGGGGATGACAAAGCTTGGCTTCTCAGAGCACCC	1499
QY	1467	TGATTTCTGTTCTGACAGAGACCCAGCTTGGCTTTAAGGATGGCAACAGTGCCTGTA	1526
DB	1500	TGATTTCTGTTCTGACAGAGACCCAGCTTGGCTTTAAGGATGGCAACAGTGCCTGTA	1559
QY	1527	TCTCCGTTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTTCTTGTGTACAAAAC	1586
DB	1560	TCTCCGTTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTTCTTGTGTACAAAAC	1619
QY	1587	ACAAGGAATAACAAACCAATAGATCTCTGGAAATGTTGTTGAGAAAGAGGCTGTA	1646
DB	1620	ACAAGGAATAACAAACCAATAGATCTCTGGAAATGTTGTTGAGAAAGAGGCTGTA	1679
QY	1647	GTGTTCTTCTCAACCGTGAAGCCGCTTCTTCCCGGAAACACCGAAAGGATCCGC	1706
DB	1680	GTGTTCTTCTCAACCGTGAAGCCGCTTCTTCCCGGAAACACCGAAAGGATCCGC	1739
QY	1707	TACTCAAAACCAAGAAATTTAAAGGAGTCTTCTTAAATTTTCAACCTTCTTCTGAGCTCA	1766
DB	1740	TACTCAAAACCAAGAAATTTAAAGGAGTCTTCTTAAATTTTCAACCTTCTTCTGAGCTCA	1799



```

QY 1767 CTTTCAGTGCATGATGATGAGATGCTGAGAGTGGCTATTAACTTTTTCCTAAG 1826
DB 1800 CTTTCAGTGCATGATGATGAGATGCTGAGAGTGGCTATTAACTTTTTCCTAAG 1859
QY 1827 ATTATGTTAAATAGATATGTTGGTTGGGAAAGTTGAAATTTTAAAGTAAATGCA 1886
DB 1860 ATTATGTTAAATAGATATGTTGGTTGGGAAAGTTGAAATTTTAAAGTAAATGCA 1919
QY 1887 TTTTAAAGTGGGAGAGGATATATCTGAGGAGCTTCAAGCATGTTGTGAACGTAT 1946
DB 1920 TTTTAAAGTGGGAGAGGATATATCTGAGGAGCTTCAAGCATGTTGTGAACGTAT 1979
QY 1947 AAAAGCACTTGAAGAGGCTCTTTCAATTTTAAAGTAAAGTAAAGTAAAGTAAAG 2006
DB 1980 AAAAGCACTTGAAGAGGCTCTTTCAATTTTAAAGTAAAGTAAAGTAAAGTAAAG 2039
QY 2007 GTAACCTAGTAGATAGAAACATCTGTCTCCGAGATGAGAGAGAACTACTATTGATTA 2066
DB 2040 GTAACCTAGTAGATAGAAACATCTGTCTCCGAGATGAGAGAGAACTACTATTGATTA 2099
QY 2067 GAGCTTAACCCAGGTTAACTGCAAGAGGCGGATCTTCAAGCTTTCCATGTAATCTG 2126
DB 2100 GAGCTTAACCCAGGTTAACTGCAAGAGGCGGATCTTCAAGCTTTCCATGTAATCTG 2159
QY 2127 TATGCAATAAGCAATGTAGTCAAGTCTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2186
DB 2160 TATGCAATAAGCAATGTAGTCAAGTCTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2219
QY 2187 CTTCATAATACACATCATGATCTCTGATGAAACATTAACAGGCCCAAGCTGTGATGTA 2246
DB 2220 CTTCATAATACACATCATGATCTCTGATGAAACATTAACAGGCCCAAGCTGTGATGTA 2279
QY 2247 TGTGACACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2306
DB 2280 TGTGACACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2339
QY 2307 GTGACAACTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2366
DB 2340 GTGACAACTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2399
QY 2367 TGGACATTTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2426
DB 2400 TGGACATTTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2459
QY 2427 TATTTCCAAATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2486
DB 2460 TATTTCCAAATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2518
QY 2487 AAAAATTTGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2546
DB 2519 AAAAATTTGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2578
QY 2547 GAACATAAACAATTTACTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2606
DB 2579 TAACATAAACAATTTACTATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2638
QY 2607 -AAAAATGAGGAGAGACAAAAA 2636
DB 2639 TAAAGTGTGGAGAGAGACAAAAA 2669

```

## RESULT 5

```

US-10-342-887-779
; Sequence 779, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.

```

```

; APPLICANT: Bernarde, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO: 779
; LENGTH: 2669
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-342-887-779

Query Match      95.7%; Score 2547.8; DB 13; Length 2669;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2591; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

QY 28 AACCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 87
DB 60 ACCTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 119
QY 88 TTCTGCTCCTGGCTGCAAGATTGCACTTATGATGATGATGATGATGATGATGATGATGATG 147
DB 120 TTCTGCTCCTGGCTGCAAGATTGCACTTATGATGATGATGATGATGATGATGATGATGATG 179
QY 148 GGAATGAAAGACCTTCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 207
DB 180 GGAATGAAAGACCTTCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 239
QY 208 ATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 267
DB 240 ATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 299
QY 268 AAAAATCTGGAAGGAGGCGGTGTGCAAGGCGGTCTGCAAGTGAATGCAAGGCGGTCTG 327
DB 300 AAAAATCTGGAAGGAGGCGGTGTGCAAGGCGGTCTGCAAGTGAATGCAAGGCGGTCTG 359
QY 328 TGGGCTCAATATTAACATTTGGGCTGAACCTGATATTTCTTATGATGATGATGATGATG 387
DB 360 TGGGCTCAATATTAACATTTGGGCTGAACCTGATATTTCTTATGATGATGATGATGATG 419
QY 388 CCAATGCGAACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 447
DB 420 CCAATGCGAACATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 479
QY 448 CATATGTTTAAACAATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 507
DB 480 CATATGTTTAAACAATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 539
QY 508 AAAAGCATCAATACGTTCTTCCCTGATGGAACCTTTTCTTCCACACACCCCGATGAGAA 567
DB 540 AAAAGCATCAATACGTTCTTCCCTGATGGAACCTTTTCTTCCACACACCCCGATGAGAA 599
QY 568 GATGGAATTTCAATCAAGCTTCTTCCACACACCTTGGTCAATTTTCCAGAAATTTGGAGATG 627
DB 600 GATGGAATTTCAATCAAGCTTCTTCCACACACCTTGGTCAATTTTCCAGAAATTTGGAGATG 659
QY 628 GTTCAGTGAAGATTTCTGTGAAACACAGGCAATGATGATGATGATGATGATGATGATGATG 687
DB 660 GTTCAGTGAAGATTTCTGTGAAACACAGGCAATGATGATGATGATGATGATGATGATGATG 719
QY 688 TGACTGTCTTACAGAGACATGACGCGGATATGTTCCATGCGCAATAGTAAAGATGCTG 747
DB 720 TGACTGTCTTACAGAGACATGACGCGGATATGTTCCATGCGCAATAGTAAAGATGCTG 779
QY 748 ACGTGTAAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 807
DB 780 ACGTGTAAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 839

```

QY 808 CATCCGAGGAACCTTCC-CAAGATCTCCCATTTATGTTGATGTCTGATTCATGATTC 866  
DB 840 CATCCGAGGAACCTTCCCAAGATCTCCCATTTATGTTGATGTCTGATTCATGATTC 899  
QY 867 CTAGCCACTTCTCATTTATTTTACACCTTAACACAGTGGAGCTTCGGGGATTAATCTG 926  
DB 900 CTAGCCACTTCTCATTTATTTTACACCTTAACACAGTGGAGCTTCGGGGATTAATCTG 959  
QY 927 GCTGTTTGTTCACCAATCATCTGTGATCAACGATGTGCTCAATGGAACCTTCA 986  
DB 960 GCTGTTTGTTCACCAATCATCTGTGATCAACGATGTGCTCAATGGAACCTTCA 1019  
QY 987 GCTTACCTCACTGGAAGGCTGAGGACCAAGACCTTGTCCGCAACGCGCAACCAAC 1046  
DB 1020 GCTTACCTCACTGGAAGGCTGAGGACCAAGACCTTGTCCGCAACGCGCAACCAAC 1079  
QY 1047 CCAGACCTTCAAAACCCCTCTTCTTATGAGACCTGCTGTGACAAACCCCTGAGCTGA 1106  
DB 1080 CCAGACCTTCAAAACCCCTCTTCTTATGAGACCTGCTGTGACAAACCCCTGAGCTGA 1139  
QY 1107 GTAGATTTCTGATGAAAACCTGCAATTAACAGATGTGCACTTTCAAGCCATCA 1166  
DB 1140 GTAGATTTCTGATGAAAACCTGCAATTAACAGATGTGCACTTTCAAGCCATCA 1199  
QY 1167 CAATGTAGAGGGAATCTTAGAGTTAATCATCATCAAGATGACAGAGCTGTGATGCCG 1226  
DB 1200 CAATGTAGAGGGAATCTTAGAGTTAATCATCATCAAGATGACAGAGCTGTGATGCCG 1259  
QY 1227 TGCCATGAGCTGAAAGCTCCCTAATAGACTTTGTCTGACCTGCAAGGAGCATTCGA 1286  
DB 1260 TGCCATGAGCTGAAAGCTCCCTAATAGACTTTGTCTGACCTGCAAGGAGCATTCGA 1319  
QY 1287 CGAGGTCTGTACATCATTTCTGACCCCACTGCAAGATCAACCAAGAGCTGTGCA 1346  
DB 1320 CGAGGTCTGTACATCATTTCTGACCCCACTGCAAGATCAACCAAGAGCTGTGCA 1379  
QY 1347 GGCCTGTGATGTGATGTGATGTGTCTGACCTGCAAGAGCTGTGATGCCG 1406  
DB 1380 GGCCTGTGATGTGATGTGATGTGTCTGACCTGCAAGAGCTGTGATGCCG 1439  
QY 1407 GGAAGTGTGTGAAAGCTTCAACCTGAGGAGATGACAAAGCTGTGCTGACGAGACC 1466  
DB 1440 GGAAGTGTGTGAAAGCTTCAACCTGAGGAGATGACAAAGCTGTGCTGACGAGACC 1499  
QY 1467 TGAATTTCTGCTGACAGAGACCCAGCTGTGCTTTAAGATGCAACAGTCCCTGA 1526  
DB 1500 TGAATTTCTGCTGACAGAGACCCAGCTGTGCTTTAAGATGCAACAGTCCCTGA 1559  
QY 1527 TCTCCGTTGCTGCTGAGCAATTTTGTCACTGTGATCTCCCTGTGCTGACAAAGAAC 1586  
DB 1560 TCTCCGTTGCTGCTGAGCAATTTTGTCACTGTGATCTCCCTGTGCTGACAAAGAAC 1619  
QY 1587 ACAAGGAATACAAACCAATAGAAATAGTCTGAGATGTGTGAGAGCAAGGCTGTA 1646  
DB 1620 ACAAGGAATACAAACCAATAGAAATAGTCTGAGATGTGTGAGAGCAAGGCTGTA 1679  
QY 1647 GTGTCTTTCTCAACCGTGCAAAAGCGGTCTTCTCCGGGAAACCAAGGAAAGATCCGC 1706  
DB 1680 GTGTCTTTCTCAACCGTGCAAAAGCGGTCTTCTCCGGGAAACCAAGGAAAGATCCGC 1739  
QY 1707 TACTCAAAAACCAAGATTTAAAGAGATTTCTTAATTTGACCTGTTCTGAGCTCA 1766  
DB 1740 TACTCAAAAACCAAGATTTAAAGAGATTTCTTAATTTGACCTGTTCTGAGCTCA 1799  
QY 1767 CTTTTCAGTGCATTTGATGTGATGTGTGAGTGTGCTTAACCTTTTCTTAAG 1826  
DB 1800 CTTTTCAGTGCATTTGATGTGATGTGTGAGTGTGCTTAACCTTTTCTTAAG 1859  
QY 1827 ATTAATTTTAAATAGATATTGTGCTTTGGGAAAGTTGAAATTTTATAGTTAAAGTCA 1886  
DB 1860 ATTAATTTTAAATAGATATTGTGCTTTGGGAAAGTTGAAATTTTATAGTTAAAGTCA 1919

QY 1887 TTTTAGAGATGGGAGGAGGATTTATCTGACGAGCTTCAGCATGTTGTGAACTGAT 1946  
DB 1920 TTTTAGAGATGGGAGGAGGATTTATCTGACGAGCTTCAGCATGTTGTGAACTGAT 1979  
QY 1947 AAAAGCACTTAGCAAGGCTTCTTTCAATTTTATTTTATGTTTCACTTAATAAGCTTAG 2006  
DB 1980 AAAAGCACTTAGCAAGGCTTCTTTCAATTTTATTTTATGTTTCACTTAATAAGCTTAG 2039  
QY 2007 GTTACTAGTATGATGAAACACTGTGTCTCCGAGAGTAAAGAGAGAGCTATTTATTA 2066  
DB 2040 GTTACTAGTATGATGAAACACTGTGTCTCCGAGAGTAAAGAGAGAGCTATTTATTA 2099  
QY 2067 GAGCTTAACCCAGTATGATGCAAGAGAGGAGGATCTTCACTTCACTGATGATG 2126  
DB 2100 GAGCTTAACCCAGTATGATGCAAGAGAGGAGGATCTTCACTTCACTGATGATG 2159  
QY 2127 TATGATTAAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2186  
DB 2160 TATGATTAAGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2219  
QY 2187 CTTCAATACACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2246  
DB 2220 CTTCAATACACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2279  
QY 2247 TGTGACACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2306  
DB 2280 TGTGACACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2339  
QY 2307 GTGACAACTTCTTCTTCTGCTGATGATGATGATGATGATGATGATGATGATGATG 2366  
DB 2340 GTGACAACTTCTTCTTCTGCTGATGATGATGATGATGATGATGATGATGATGATG 2399  
QY 2367 TGGACATTTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2426  
DB 2400 TGGACATTTAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2459  
QY 2427 TATTTCAAAATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 2486  
DB 2460 TATTTCAAAATTTTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 2518  
QY 2487 AAAAATTTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2546  
DB 2519 AAAAGATGAGTCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2578  
QY 2547 GAATTAACCAATTTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 2606  
DB 2579 TAACTAAACCAATTTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 2638  
QY 2607 -AAAATGAGGAGAGACAAAAA 2636  
DB 2639 TAAAGTGTGGAAGAGACAAAAA 2669

RESULT 6  
US-10-172-118-779  
; Sequence 779, Application US/10172118  
; Publication No. US20030224374A1  
; GENERAL INFORMATION:  
; APPLICANT: Dai, HongYue  
; APPLICANT: He, Yudong  
; APPLICANT: Linsley, Peter  
; APPLICANT: Mao, Mao  
; APPLICANT: Roberts, Chris  
; APPLICANT: Van 't Veer, Marc  
; APPLICANT: Van de Vijver, Marc  
; APPLICANT: Bernards, Rene  
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients  
; FILE REFERENCE: 9301-175-999  
; CURRENT APPLICATION NUMBER: US/10/172, 118  
; PRIOR FILING DATE: 2002-06-14  
; PRIOR APPLICATION NUMBER: 60/380,770  
; PRIOR FILING DATE: 2002-05-14  
; NUMBER OF SEQ ID NOS: 2699





Db 780 AGGTGTACAGATCGATTCCTGTTGTCATATGTTCCAGAAAGACATCGAATT 839  
 Qy 808 CATCCGAGCAAACTTCC - CAAAGATCTCCCATTAATGTTGATGTCTGATTCATGATC 866  
 Db 840 CATCCGAGCAAACTTCC CAAAGATCTCCCATTAATGTTGATGTCTGATTCATGATC 899  
 Qy 867 CTAGCCACTTCTCAATTATTTCTACATTAACTACAGGTGAGCTCGGGATTAATAGTG 926  
 Db 900 CTAGCCACTTCTCAATTATTTCTACATTAACTACAGGTGAGCTCGGGATTAATAGTG 959  
 Qy 927 GCCTGTTGTTCCACCATCATCTGTAATCAGATGATGCTCAATGGAACCTTCA 986  
 Db 960 GCCTGTTGTTCCACCATCATCTGTAATCAGATGATGCTCAATGGAACCTTCA 1019  
 Qy 987 GCCTTAACCTCACTGTGAAGCTGACGACCAAGACCTTGTCCGACACCGCACACAC 1046  
 Db 1020 GCCTTAACCTCACTGTGAAGCTGACGACCAAGACCTTGTCCGACACCGCACACAC 1079  
 Qy 1047 CCAGACCTTCAAAACCCCTCTTTTAGACCTGTGTGTACAACCCCTGTGAGCTGA 1106  
 Db 1080 CCAGACCTTCAAAACCCCTCTTTTAGACCTGTGTGTACAACCCCTGTGAGCTGA 1139  
 Qy 1107 GTAGGATTCGATGAAACCTCCAGATTACAGATGAGCCACTTTGAAGCCACATCA 1166  
 Db 1140 GTAGGATTCGATGAAACCTCCAGATTACAGATGAGCCACTTTGAAGCCACATCA 1199  
 Qy 1167 CAATTGTAGAGGAACTTTAGAGGTTAACATCATCCAGATGACAGACCTCTGATGCCG 1226  
 Db 1200 CAATTGTAGAGGAACTTTAGAGGTTAACATCATCCAGATGAGCCACTTTGAGCCG 1259  
 Qy 1227 TGGCATGGCTGAAAGCTCCCTTAATGACTTTGTGTGTGACCTGCGCAAGGAGCATTC 1286  
 Db 1260 TGGCATGGCTGAAAGCTCCCTTAATGACTTTGTGTGTGACCTGCGCAAGGAGCATTC 1319  
 Qy 1287 CGAGAGTCTGTACATTCATTTCTGACCCGACCTGGAGATCACCAAGACAGTGTGA 1346  
 Db 1320 CGAGAGTCTGTACATTCATTTCTGACCCGACCTGGAGATCACCAAGACAGTGTGA 1379  
 Qy 1347 GCCCTGTGATGTGATGATGTGTCTGCTGACTGTGAGACGAACCTTCAATGGAGTCTG 1406  
 Db 1380 GCCCTGTGATGTGATGATGTGTCTGCTGACTGTGAGACGAACCTTCAATGGAGTCTG 1439  
 Qy 1407 GGAAGTACTGTGTGAACCTCACCTGGGGGATGACCAAGGCTGGCTCTCAGAGACACC 1466  
 Db 1440 GGAAGTACTGTGTGAACCTCACCTGGGGGATGACCAAGGCTGGCTCTCAGAGACACC 1499  
 Qy 1467 TGAATTTCTGTCTGACAGAGACCAAGCTGGCTTTAGATGAGTGAACAGTGGCCCTGA 1526  
 Db 1500 TGAATTTCTGTCTGACAGAGACCAAGCTGGCTTTAGATGAGTGAACAGTGGCCCTGA 1559  
 Qy 1527 TCTCGTTGGCTGTGTGGCCATATTTTGTCACTGTGATCTCCCTTGTGTGTAACAAAAC 1586  
 Db 1560 TCTCGTTGGCTGTGTGGCCATATTTTGTCACTGTGATCTCCCTTGTGTGTAACAAAAC 1619  
 Qy 1587 ACAAGGAATACACCCATATGAAATAGTCTGGGAGATGTGTGACAGACAAAGGCTTGA 1646  
 Db 1620 ACAAGGAATACACCCATATGAAATAGTCTGGGAGATGTGTGACAGACAAAGGCTTGA 1679  
 Qy 1647 GGTGCTTTCTCAACGCTGCAAAAGCCGTGTCTTCCCGGGAACCAAGAAAGAGTCCGC 1706  
 Db 1680 GGTGCTTTCTCAACGCTGCAAAAGCCGTGTCTTCCCGGGAACCAAGAAAGAGTCCGC 1739  
 Qy 1707 TACTCAAAAACCAAGAAATTTAAAGGATTTCTTAATTTGACCTTGTGTAAGCTGA 1766  
 Db 1740 TACTCAAAAACCAAGAAATTTAAAGGATTTCTTAATTTGACCTTGTGTAAGCTGA 1799  
 Qy 1767 CTTTTCAGTGCATTGATGTGAGTGTGTGAGTGTGATTAACCTTTTTCCTAAG 1826  
 Db 1800 CTTTTCAGTGCATTGATGTGAGTGTGTGAGTGTGATTAACCTTTTTCCTAAG 1859  
 Qy 1827 ATTATTTAAADAGATTTGTGTTGGGAAAGTTGAATTTTATAGGTTAAATGTCA 1886

Db 1860 ATTATTTAAATAGATTTGTGTTGGGGAAGTGAATTTTATAGTAAATGTCA 1919  
 Qy 1887 TTTTGAAGATGGGAGAGAGATTTATACGACAGGACCTTCAAGCATTTGTGAAAATGAT 1946  
 Db 1920 TTTTGAAGATGGGAGAGAGATTTATACGACAGGACCTTCAAGCATTTGTGAAAATGAT 1979  
 Qy 1947 AAAAGCAATTAGCAAGGCTTTTCAATTATTTTATGTTTCACTTATTAAGTCTTG 2006  
 Db 1980 AAAAGCAATTAGCAAGGCTTTTCAATTATTTTATGTTTCACTTATTAAGTCTTG 2039  
 Qy 2007 GTAACTAGAGATGAAACACTGTGTCCCGAGATTAAGGAGAGAGACTATGATTA 2066  
 Db 2040 GTAACTAGAGATGAAACACTGTGTCCCGAGATTAAGGAGAGAGACTATGATTA 2099  
 Qy 2067 GAGCCTAACCCAGGTTTAACTGCAAGAGAGGCGGATCTTCAAGCTTCCATGTAAGT 2126  
 Db 2100 GAGCCTAACCCAGGTTTAACTGCAAGAGAGGCGGATCTTCAAGCTTCCATGTAAGT 2159  
 Qy 2127 TATGATTAAGCCAAATGTAGTCAAGTTTCTAAGATATGTTCGAAGCTAATGCCA 2186  
 Db 2160 TATGATTAAGCCAAATGTAGTCAAGTTTCTAAGATATGTTCGAAGCTAATGCCA 2219  
 Qy 2187 CTTCAATACACATCATGAACTCCGATGAGAACATTAACAGGCCCAAGCTGAGATGA 2246  
 Db 2220 CTTCAATACACATCATGAACTCCGATGAGAACATTAACAGGCCCAAGCTGAGATGA 2279  
 Qy 2247 TGTGACACTTGTGACTGACGAAAAAAATACCTCTCATTAATGGGTGGAGATTTTG 2306  
 Db 2280 TGTGACACTTGTGACTGACGAAAAAAATACCTCTCATTAATGGGTGGAGATTTTG 2339  
 Qy 2307 GTGACAACTTACTTTGCTTGTGAGTGAAGAAATGATTAATATATTCATTATTCGA 2366  
 Db 2340 GTGACAACTTACTTTGCTTGTGAGTGAAGAAATGATTAATATATTCATTATTCGA 2399  
 Qy 2367 TGGACATTAAGTATGCTTTTATATACAGGCAATGATGTGAGNACACTTGTGTA 2426  
 Db 2400 TGGACATTAAGTATGCTTTTATATACAGGCAATGATGTGAGNACACTTGTGTA 2459  
 Qy 2427 TATTTCCAAATTTTGTATAGTGTGACATATTTTGAATCAAAATATTAAGACTTTTC 2486  
 Db 2460 TATTTCCAAATTTTGTATAGTGTGACATATTTTGAATC - AATATTAAGACTTTTC 2518  
 Qy 2487 AAAAATTTGTTGCTTGTGTTTTCATGCAACTTATGATGATGATTTCCCTGTGTTG 2546  
 Db 2519 AAAAATTTGTTGCTTGTGTTTTCATGCAACTTATGATGATGATTTCACTGTGTTG 2578  
 Qy 2547 GAACATAAACCACTTACTATATGTTAGACAAGACATTTTCTTCTCTGTAAGAA 2606  
 Db 2579 TACTTAATAACCACTTACTATATGTTAGACATGACATTTCTTCTCTCTGTAAGAA 2638  
 Qy 2607 - AAAATGAGGAGAGACAAAAA 2636  
 Db 2639 TAAAGTGTGGAGAGACAAAAA 2669  
 RESULT 8  
 US-10-450-826-92  
 ; Sequence 92, Application US/10450826  
 ; Publication No. US20040101818A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ji, Darren  
 ; APPLICANT: Axelrod, Douglas W.  
 ; APPLICANT: Cook, Jonathan S.  
 ; APPLICANT: Jaiswal, Neelam  
 ; APPLICANT: Bistefni, Richard  
 ; APPLICANT: Houghton, Adam  
 ; APPLICANT: Mettz, Lawrence  
 ; TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiation  
 ; FILE REFERENCE: 044921-5039-WO  
 ; CURRENT APPLICATION NUMBER: US/10/450,826  
 ; CURRENT FILING DATE: 2003-06-18  
 ; PRIOR APPLICATION NUMBER: US 60/255,882  
 ; PRIOR FILING DATE: 2000-12-18



```

Db 1980 AAAAGCACTTAGCAAGGCTCTTTTCATTAATTTTATGTTTCACTTAATAAGCTTAG 2039
QY 2007 GTAACAGTAGAGTAGAAACACTGTGTCCCGAGTAAGAGAGAAGCTACTATTGATTA 2066
Db 2040 GTAACTAGTAGATAGAAACCTGTGTCCCGAGTAGTAAGAGAGAAGCTACTATTGATTA 2099
QY 2067 GAGCTTAACCCAGGTAACTGCAAGAGAGGGGAGTACTTTCAGCTTTCATGTACTG 2126
Db 2100 GAGCTTAACCCAGGTAACTGCAAGAGAGGGGAGTACTTTCAGCTTTCATGTACTG 2159
QY 2127 TATGCAATAAGCAATGTAGTCCAGTTTCAAGATATATGTTCCAACTAAGTAATCCCA 2186
Db 2160 TAGCAATAAGCAATGTAGTCCAGTTTCAAGATATATGTTCCAACTAAGTAATCCCA 2219
QY 2187 CTTCATATACACACTGAGTAAGTCCAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 2246
Db 2220 CTTCATATACACACTGAGTAAGTCCAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 2279
QY 2247 TGTGCACTGTGTAGACTCAGAAAAATACTACTCAATAATGAGTGGAGTATTTTG 2306
Db 2280 TGTGCACTGTGTAGACTCAGAAAAATACTACTCAATAATGAGTGGAGTATTTTG 2339
QY 2307 GTGACAACTGACTGCTGTGGCTGAGTGAAGAAATGATATATCAATATCAATTTATCCA 2366
Db 2340 GTGACAACTGACTGCTGTGGCTGAGTGAAGAAATGATATATCAATATCAATTTATCCA 2399
QY 2367 TGGACATTTAGTAGTGTCTTTTATATACAGGATGATGCTGAGTGAAGTCTTGTGTA 2426
Db 2400 TGGACATTTAGTAGTGTCTTTTATATACAGGATGATGCTGAGTGAAGTCTTGTGTA 2459
QY 2427 TATTTCCAAATTTTGTATATAGTGTGACATATTTGAAATCAAAATATTAGACTTTCC 2486
Db 2460 TATTTCCAAATTTTGTATAGTGTGACATATTTGAAATC-ATAATATTAAGACTTTCC 2518
QY 2487 AAAAATTTGTCCTGCTGTTTTCATGAGCACTTGATCAGTAAGAGATTTCCCTCTGTTG 2546
Db 2519 AAAGATGAGGTCCGTGTGTTTTCATGAGCACTTGATCAGTAAGATTTCCCTCTGTTG 2578
QY 2547 GAACATAAACCATTTACTATATGTAGACAGACATTTTTCCTTCCCTGAAAAA 2606
Db 2579 TAACATAAACCATTTACTATATGTAGACATGACATTTTTCCTCTCTGAAAAA 2638
QY 2607 -AAATGAGGAAAGACAAAAA 2636
Db 2639 TAAAGTGTGGAGAGACAAAAA 2669

RESULT 9
US-10-084-817-169
; Sequence 169, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Ued G. Nuchtern
; APPLICANT: Sharon E. Pion
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084,817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270,784
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 169
; LENGTH: 2666
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 2687977GB1
US-10-084-817-169

```

```

Query Match 95.1%; Score 2531.6; DB 15; Length 2666;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2576; Conservative 0; Mismatches 19; Indels 3; Gaps 3;

QY 28 AACCTTGTGCTGTGCTCGTGCATGAGTAATGATGATGATGATGATGATGATGATGATGATGAT 87
Db 70 ACCCTTGAAGTGTGCTGTGCTCGTGCATGAGTAATGATGATGATGATGATGATGATGATGATGATGAT 129
QY 88 TTCTGCTCTGTGCTGTGCTCGTGCATGAGTAATGATGATGATGATGATGATGATGATGATGATGAT 147
Db 130 TTCTGCTCTGTGCTGTGCTCGTGCATGAGTAATGATGATGATGATGATGATGATGATGATGATGAT 189
QY 148 GCATGAAAGACCTTGTGCTGTGATGAGGAGACATATATTAATGATGATGATGATGATGATGATGATGAT 207
Db 190 GCATGAAAGACCTTGTGCTGTGATGAGGAGACATATATTAATGATGATGATGATGATGATGATGATGAT 249
QY 208 ATGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 267
Db 250 ATGAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 309
QY 268 AAAAAGCTGTGAAAGGAGGCGGTGTGACAGGCGGTCTGAACAATGATCTACAGAGCCCTCG 327
Db 310 AAAAAGCTGTGAAAGGAGGCGGTGTGACAGGCGGTCTGAACAATGATCTACAGAGCCCTCG 369
QY 328 TGGGCTCAATATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 387
Db 370 TGGGCTCAATATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 429
QY 388 CCAATGAGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 447
Db 430 CCAATGAGCAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 489
QY 448 CATATGTTTACAACTGAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 507
Db 490 CATATGTTTACAACTGAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 549
QY 508 AAAGCATATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 567
Db 550 AAAGCATATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 609
QY 568 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 627
Db 610 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 669
QY 628 GTTCACTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
Db 670 GTTCACTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 729
QY 688 TGAATGCTTACAGAAACATGAGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 747
Db 730 TGAATGCTTACAGAAACATGAGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
QY 748 AGGTGATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 807
Db 790 AGGTGATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849
QY 808 CATCCAGCAAAACCTTCC-CAAGATCTCCCATTAATGATGATGATGATGATGATGATGATGATGATGAT 866
Db 850 CATCCAGCAAAACCTTCCCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 909
QY 867 CTAGCCACTTCCCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 926
Db 910 CTAGCCACTTCCCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 969
QY 927 GCCTGTTGTTTCCCAATCACTAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 986
Db 970 GCCTGTTGTTTCCCAATCACTAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1029
QY 987 GCCTTAACCTCACTGTAAGAGCTGACAGACAGAGACCTTGTCCGCAACCCGCAACACAC 1046
Db 1030 GCCTTAACCTCACTGTAAGAGCTGACAGACAGAGACCTTGTCCGCAACCCGCAACACAC 1089
QY 1047 CCAAGCTTCAAAACCACCCCTTTTATGAGACTGCTGTGTGACAAACCCCTGAGAGCTGA 1106

```



Db 1090 CCAGACCTTCAAAACCCACCCCTTCTTAGAGACCTGGTGGTGAACAACCCCTGGAGCTGA 1149  
Qy 1107 GTAGATTCTCGTAGAAAACTGCCAGATTACAGATATGGCACTTTTAACCCACATCA 1166  
Db 1150 GTAGATTCTCGTAGAAAACTGCCAGATTACAGATATGGCACTTTTAACCCACATCA 1209  
Qy 1167 CAATTGTAGAGGAACTTTAGAGTTAAATCATCATCCAGATGACAGACGTCCTGATGCCG 1226  
Db 1210 CAATTGTAGAGGAACTTTAGAGTTAAATCATCATCCAGATGACAGACGTCCTGATGCCG 1269  
Qy 1227 TGCCATGGCCCTGAAAGCTCCCTAATAGACTTTGTGTGACCTGCCAAGAGACATTCGA 1286  
Db 1270 TGCCATGGCCCTGAAAGCTCCCTAATAGACTTTGTGTGACCTGCCAAGAGACATTCGA 1329  
Qy 1287 CGAGAGTCTGTACATCATTTCTGACCCCACTGCGAGATCCCAACCAAGACATTCGCA 1346  
Db 1330 CGAGAGTCTGTACATCATTTCTGACCCCACTGCGAGATCCCAACCAAGACATTCGCA 1389  
Qy 1347 GCCCTGTGATGTGATGAGATGTGTCTGCTGACTGTAGAGCAACCTTCAATGGATCTG 1406  
Db 1390 GCCCTGTGATGTGATGAGATGTGTCTGCTGACTGTAGAGCAACCTTCAATGGATCTG 1449  
Qy 1407 GGAGGTACTGTGTAACTCACTCGGAGGATGACAAAGCTGCTCTCAAGCAACC 1466  
Db 1450 GGAGGTACTGTGTAACTCACTCGGAGGATGACAAAGCTGCTCTCAAGCAACC 1509  
Qy 1467 TGATTGTGTTCCCGAGAGAGACCAGCTCGCTTTAAGATGGCAAAAGTCCCTGA 1526  
Db 1510 TGATTGTGTTCCCGAGAGAGACCAGCTCGCTTTAAGATGGCAAAAGTCCCTGA 1569  
Qy 1527 TCTCCGTGGCTGTGGCCATATTTGTCACTGTGATCTCCCTTGTGTGACAAAAC 1586  
Db 1570 TCTCCGTGGCTGTGGCCATATTTGTCACTGTGATCTCCCTTGTGTGACAAAAC 1629  
Qy 1587 ACAAGGATACACCCCAATGAAATATGCTCTGGAAATGTGTGACAAAGGCTGA 1646  
Db 1630 ACAAGGATACACCCCAATGAAATATGCTCTGGAAATGTGTGACAAAGGCTGA 1689  
Qy 1647 GTGCTCTTCTCAACCGTGAACCGCTGTCTTCCGGGAAACCAAGAAAGATCCGC 1706  
Db 1690 GTGCTCTTCTCAACCGTGAACCGCTGTCTTCCGGGAAACCAAGAAAGATCCGC 1749  
Qy 1707 TACTCAAAAACCAAGATTTTAAAGAGATTCTTAAATTTGACCTGTTCTGAAGCTCA 1766  
Db 1750 TACTCAAAAACCAAGATTTTAAAGAGATTCTTAAATTTGACCTGTTCTGAAGCTCA 1809  
Qy 1767 CTTTTCAGTCCATGATGTGAGATGTGCTGAGTGGCTATTAACCTTTTTTCTTAAG 1826  
Db 1810 CTTTTCAGTCCATGATGTGAGATGTGCTGAGTGGCTATTAACCTTTTTTCTTAAG 1869  
Qy 1827 ATTATGTTAAATAGATATGTTGGGAAAGTTGAATTTTAAAGTTAAAGTCA 1886  
Db 1870 ATTATGTTAAATAGATATGTTGGGAAAGTTGAATTTTAAAGTTAAAGTCA 1929  
Qy 1887 TTTTAAAGATGGGAGAGGATTAATCTGAGGACAGCTTCAGCATGTTGAAAAGTAT 1946  
Db 1930 TTTTAAAGATGGGAGAGGATTAATCTGAGGACAGCTTCAGCATGTTGAAAAGTAT 1989  
Qy 1947 AAAAGCACTTACCAAGGCTTTCTTTTCATTAATTTTATGTTCACTTAAGTCTTAG 2006  
Db 1990 AAAAGCACTTACCAAGGCTTTCTTTTCATTAATTTTATGTTCACTTAAGTCTTAG 2049  
Qy 2007 GTAAGTAGTAGAGAAACACTGTGTCCGAGAGTAGAGAGAGAGTAGTAGTATTA 2066  
Db 2050 GTAAGTAGTAGAGAAACACTGTGTCCGAGAGTAGAGAGAGAGTAGTAGTATTA 2109  
Qy 2067 GAGCTTAACCAAGGTTAACTGCAAGAGAGGCGGATACCTTTAGCTTTCACTGATAG 2126  
Db 2110 GAGCTTAACCAAGGTTAACTGCAAGAGAGGCGGATACCTTTAGCTTTCACTGATAG 2169  
Qy 2127 TATGCAATAAGCCAAATGTAGTCAAGTTCTTAAGATCATGTTCCAACTTAAGTATCCA 2186

Db 2170 TATGCAATAAGCCAAATGTAGTCCAGTTCTTAAGATCATGTTCCAACTTAAGTATCCA 2229  
Qy 2187 CTTCAATACACACTCATGAACTCTCTGATGGAAACAATAACAGGCCCAAGCTGTGGTATGA 2246  
Db 2230 CTTCAATACACACTCATGAACTCTCTGATGGAAACAATAACAGGCCCAAGCTGTGGTATGA 2289  
Qy 2247 TGTGCACTTGTGAGCTGAGAAAAAATCTACTCATTAATAGGGGAGATATTTTG 2306  
Db 2290 TGTGCACTTGTGAGCTGAGAAAAAATCTACTCATTAATAGGGGAGATATTTTG 2349  
Qy 2307 GTGCAACCTACTTGTGCTGAGTGAAGAGATGATATATCATATTAATTAATTAAGCTTCC 2366  
Db 2350 GTGCAACCTACTTGTGCTGAGTGAAGAGATGATATATCATATTAATTAATTAAGCTTCC 2409  
Qy 2367 TGAACATTTAGTATGCTTTTATTAATACAGGATATGCTGAGTACACTCTTGTGTA 2426  
Db 2410 TGAACATTTAGTATGCTTTTATTAATACAGGATATGCTGAGTACACTCTTGTGTA 2469  
Qy 2427 TATTTCCAAATTTTGTATGATGCTGACATATTTGAATCAAAATTAAGCTTCC 2486  
Db 2470 TATTTCCAAATTTTGTATGATGCTGACATATTTGAATCAAAATTAAGCTTCC 2528  
Qy 2487 AAAAATTTGCTGCTGCTTTTTCATGCAACTGATCAGTAAGATTTCCCTGTGTTG 2546  
Db 2529 AAAGATAGGCTCCCTGCTTTTTCATGCAACTGATCAGTAAGATTTCCCTGTGTTG 2588  
Qy 2547 GAACTAAACCACTTACTATATGTAGACAAAGCAATTTTTTTTTTCTCTCGAAAAA 2606  
Db 2589 TAACTAAACCACTTACTATATGTAGACAAAGCAATTTTTTTTTTCTCTCGAAAAA 2648  
Qy 2607 -AAAATGAGGAGAGAGC 2623  
Db 2649 TAAAGTGTGGAGAGAGAC 2666

RESULT 10  
US-10-219-535-41  
; Sequence 41. Application US/10219535  
; Publication No. US20040404179A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Geriltsen, Mary  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Guiney, Austin L.  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stephan, Jean-Philippe F.  
; APPLICANT: Matanabe, Colin L.  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE OF INVENTION: ACIDS ENCODING THE SAME  
; FILE REFERENCE: P3530P1C60  
; CURRENT APPLICATION NUMBER: US/10/219,535  
; PRIOR FILING DATE: 2002-08-14  
; PRIOR APPLICATION NUMBER: 10/119,480  
; PRIOR FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: 60/059113  
; PRIOR FILING DATE: 1997-09-17  
; PRIOR APPLICATION NUMBER: 60/062287  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/063549  
; PRIOR FILING DATE: 1997-10-28  
; PRIOR APPLICATION NUMBER: 60/064103  
; PRIOR FILING DATE: 1997-10-31  
; PRIOR APPLICATION NUMBER: 60/069873  
; PRIOR FILING DATE: 1997-12-17  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/079294  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: 60/079656





QY 1971 TTCAATATTTTATGTTCACTTATAAAGCTTAGGTAAGGATAGAAACCTG 2030  
 Db 2030 TTCAATATTTTATGTTCACTTATAAAGCTTAGGTAAGGATAGAAACCTG 2089  
 QY 2031 TGTCCCGAGAGTANGAGAGAGCTACTATTGATTAGAGCTTAACCCAGTTAACTGAA 2090  
 Db 2090 TGTCCCGAGAGTANGAGAGAGCTACTATTGATTAGAGCTTAACCCAGTTAACTGAA 2149  
 QY 2091 GAAGAGCGGAGTACTTACCTTCCATGTAATGATATGATTAAGCCAAATGATGCCA 2150  
 Db 2150 GAAGAGCGGAGTACTTACCTTCCATGTAATGATATGATTAAGCCAAATGATGCCA 2209  
 QY 2151 GTTCTTAAGATCATGTTCCAGCTAAGTAACTGATCCCACTTCAATACACACTGATGATCC 2210  
 Db 2210 GTTCTTAAGATCATGTTCCAGCTAAGTAACTGATCCCACTTCAATACACACTGATGATCC 2269  
 QY 2211 TGATGAAACAATAACAGAGCCCAAGCTGTGTGATGATGATGACACTTGTGATGACAA 2270  
 Db 2270 TGATGAAACAATAACAGAGCCCAAGCTGTGTGATGATGATGACACTTGTGATGACAA 2329  
 QY 2271 AAAATACCTACTCTCATTAATGGGTGGAGATATTTGGTGAACAACCTTCTTGTGCTG 2330  
 Db 2330 AAAATACCTACTCTCATTAATGGGTGGAGATATTTGGTGAACAACCTTCTTGTGCTG 2389  
 QY 2331 AGTGAAGAAATGATATTCATATATTCATTTATTCATGACATTTAGTATGATGCTTTTA 2390  
 Db 2390 AGTGAAGAAATGATATTCATATATTCATTTATTCATGACATTTAGTATGATGCTTTTA 2449  
 QY 2391 TATACCCGAGATGATGCTGATGACACTCTGTGTGATATTTCCAAATTTTGTATGATG 2450  
 Db 2450 TATACCCGAGATGATGCTGATGACACTCTGTGTGATATTTCCAAATTTTGTATGATG 2509  
 QY 2451 CTGACATATTTGAAATCAAAATTTAAGACTTCCAAATTTGGTCCCTGCTGTTTCA 2510  
 Db 2510 CTGACATATTTGAAATCAAAATTTAAGACTTCCAAATTTGGTCCCTGCTGTTTCA 2568  
 QY 2511 TGGCAACTGATGATGATGATTTCCCTCTGTTTGGAACTAAACCATTTACTATATGT 2570  
 Db 2569 TGGCAACTGATGATGATGATTTCCCTCTGTTTGGAACTAAACCATTTACTATATGT 2628  
 QY 2571 TAAACAAGACATTTTCTTCTCTGTAAGAA-AAAATAGAGGAGAGACA 2624  
 Db 2629 TAAACAAGACATTTTCTTCTCTGTAAGAA-AAAATAGAGGAGAGACA 2683  
 RESULT 11  
 US-10-232-230-41  
 ; Sequence 41, Application US/10232230  
 ; Publication No. US2004044180A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Geriltsen, Mary  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Stephan, Jean-Philippe F.  
 ; APPLICANT: Watanabe, Colin L.  
 ; APPLICANT: Wood, William I.  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE OF INVENTION: ACIDS ENCODING THE SAME  
 ; FILE REFERENCE: P3530P103  
 ; CURRENT APPLICATION NUMBER: US/10/232, 230  
 ; PRIOR FILING DATE: 2002-08-29  
 ; PRIOR APPLICATION NUMBER: 10/119, 480  
 ; PRIOR FILING DATE: 2002-04-09  
 ; PRIOR APPLICATION NUMBER: 60/059113  
 ; PRIOR FILING DATE: 1997-09-17  
 ; PRIOR APPLICATION NUMBER: 60/062287  
 ; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/063549  
 ; PRIOR FILING DATE: 1997-10-28  
 ; PRIOR APPLICATION NUMBER: 60/064103  
 ; PRIOR FILING DATE: 1997-10-31  
 ; PRIOR APPLICATION NUMBER: 60/069873  
 ; PRIOR FILING DATE: 1997-12-17  
 ; PRIOR APPLICATION NUMBER: 60/078910  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/079294  
 ; PRIOR FILING DATE: 1998-03-25  
 ; PRIOR APPLICATION NUMBER: 60/079656  
 ; PRIOR FILING DATE: 1998-03-26  
 ; PRIOR APPLICATION NUMBER: 60/079728  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 246  
 ; SEQ ID NO 41  
 ; LENGTH: 2683  
 ; TYPE: DNA  
 ; ORGANISM: Homo Sapien  
 US-10-232-230-41  
 Query Match 93.4%; Score 2485; DB 13; Length 2683;  
 Best Local Similarity 97.8%; Pred. No. 0;  
 Matches 2576; Conservative 0; Mismatches 20; Indels 39; Gaps 4;  
 QY 28 AACCTGTCCTGCGTCCCTGATGAAATTCAGATGAAATGCTCTACTATTTCTGGAAT 87  
 Db 50 ACCTTAGTGCCTGCGTCCCTGATGAAATTCAGATGAAATGCTCTACTATTTCTGGAAT 109  
 QY 88 TTCTGCTCCGCGTGAAGATGTCGCTGATGAGGAGACATCAATTAATGCTGCTGCTG 147  
 Db 110 TTCTGCTCCGCGTGAAGATGTCGCTGATGAGGAGACATCAATTAATGCTGCTGCTG 169  
 QY 148 GCAATGAAGACCTTGTGCTTACATGAGGAGACATCAATTAATGCTGCTGCTG 207  
 Db 170 GCAATGAAGACCTTGTGCTTACATGAGGAGACATCAATTAATGCTGCTGCTG 229  
 QY 208 ATGAATAATGCTGGAATGAAATCTTACCACTGTGGAAGCGGAGACATGAGTGA 267  
 Db 230 ATGAATAATGCTGGAATGAAATCTTACCACTGTGGAAGCGGAGACATGAGTGA 289  
 QY 268 AAAAAGCTGGAAGGAGGCGGTGTCAGGCGGTCTGACCAATGATCAACGCGCTCG 327  
 Db 290 AAAAAGCTGGAAGGAGGCGGTGTCAGGCGGTCTGACCAATGATCAACGCGCTCG 349  
 QY 328 TGGGCTCAATATATATTCATTTGCGTGAACCTGATATTTCCCTAGATCCAAAAGAGATG 387  
 Db 350 TGGGCTCAATATATATTCATTTGCGTGAACCTGATATTTCCCTAGATCCAAAAGAGATG 409  
 QY 388 CCAATGGCAACATATGCTATGAGAGAACTGCAAGAAATGAGCTGTTATCTGCTGATC 447  
 Db 410 CCAATGGCAACATATGCTATGAGAGAACTGCAAGAAATGAGCTGTTATCTGCTGATC 469  
 QY 448 CATATGTTACAACTGAGCAGATGTCAGAGACAGTACAGGCGGAAATGAGACCGGCC 507  
 Db 470 CATATGTTACAACTGAGCAGATGTCAGAGACAGTACAGGCGGAAATGAGACCGGCC 529  
 QY 508 AAAGCCATATACGCTTCTCTGATGAGGAAACCTTTCTCAACCAACCCCGATGAGAA 567  
 Db 530 AAAGCCATATACGCTTCTCTGATGAGGAAACCTTTCTCAACCAACCCCGATGAGAA 589  
 QY 568 GATGGAATTTATCTACGCTTCTCAACACTTGTGATGATTTCCAGAAATTTGGAGCAT 627  
 Db 590 GATGGAATTTATCTACGCTTCTCAACACTTGTGATGATTTCCAGAAATTTGGAGCAT 649  
 QY 628 GTTCAAGTGAAGTTCTGGAACACAGCCCAATGTCATTTGGGCTCAACTATGGAAG 687  
 Db 650 GTTCAAGTGAAGTTCTGGAACACAGCCCAATGTCATTTGGGCTCAACTATGGAAG 709  
 QY 688 TCACTGTCTTACAGAGACATGAGCGGCAATATGTTCCATGCAAGTGAAGATGTGT 747  
 Db 710 TCACTGTCTTACAGAGACATGAGCGGCAATATGTTCCATGCAAGTGAAGATGTGT 769

QY	748	ACGGGGTAAAGATGACATGCTCGGGTTTGATGCTAATGTCGAGAAAGAAAGATCGAAAT	807
Dp	770	ACGGGTAAACAGATCAGATTCCTGGTGTGTTGATGTTGATGTCAGAAAGAAAGATCGAAAT	829
QY	808	CATCCGACGAACCTTCC - CAAAGATCTCCCATATATGTTGATGTCATCATCATGATC	866
Dp	830	CATCCGACGAACCTTCTCTCAAGATCTCCCATATATGTTGATGTCATCATCATGATC	889
QY	867	CTAGCCACTTCTCAATTAATTTCTACATTAACAAGTGAAGCTTCGGGATATATCTG	926
Dp	890	CTAGCCACTTCTCAATTAATTTCTACATTAACAAGTGAAGCTTCGGGATATATCTG	949
QY	927	GCCGTGTTGTTCCACCAATCATATCTGTGATCAGACGTATGTCCTAAATGSAACCTTCA	986
Dp	950	GCCGTGTTGTTCCACCAATCATATCTGTGATCAGACGTATGTCCTAAATGSAACCTTCA	1009
QY	987	GCCCTAACCTCATCTGTGAAGCTGACACACAGAGACTTCTCCGCCACCGCCACACAC	1046
Dp	1010	GCCCTAACCTCATCTGTGAAGCTGACACACAGAGACTTCTCCGCCACCGCCACACAC	1069
QY	1047	CCAGACCTTGAAAAACCCACCCCTTCTTT-----	1074
Dp	1070	CCAGACCTTGAAAAACCCACCCCTTCTTTAGCACTACTAAATCTTATGATCAACA	1129
QY	1075	----AGACCTGCTGTGTGACACCCCTCTGAGCTGAGTAGATTCCTGATGAACCTGCC	1130
Dp	1130	CCCCAGACCTCATCTGTGTGACACCCCTCTGAGCTGAGTAGATTCCTGATGAACCTGCC	1189
QY	1131	AGATTTAACAGATATGAGCCACTTTCAAGCCACATCATCAATTGTGAGAGGAATCTTAAGG	1190
Dp	1190	AGATTTAACAGATATGAGCCACTTTCAAGCCACATCATCAATTGTGAGAGGAATCTTAAGG	1249
QY	1191	TTAACATCATCAGATGACAGAGCTCTGTGATGCGGTGCAATGAGCTTGAAAGCTCCCTAA	1250
Dp	1250	TTAACATCATCAGATGACAGAGCTCTGTGATGCGGTGCAATGAGCTTGAAAGCTCCCTAA	1309
QY	1251	TAGACTTTGTGTGATGCTGCCCAAGGAGACATTTCCACAGAGAGTCTGTACCATATTTCTG	1310
Dp	1310	TAGACTTTGTGTGATGCTGCCCAAGGAGACATTTCCACAGAGAGTCTGTACCATATTTCTG	1369
QY	1311	ACCCACCTGGGAGATGACCCAGAAACAAGCTGTGAGGCTGTGAGATGTGATGATGAT	1370
Dp	1370	ACCCACCTGGGAGATGACCCAGAAACAAGCTGTGAGGCTGTGAGATGTGATGATGAT	1429
QY	1371	GTCGTGTGACGTGTGAGCAAACTTCAATGAGTCTGGGACGTACTGTGTGAACCTCACCC	1430
Dp	1430	GTCGTGTGACGTGTGAGCAAACTTCAATGAGTCTGGGACGTACTGTGTGAACCTCACCC	1489
QY	1431	TGGGGGATGACACAAGCTGTGCTCACAGACACCTGTATTTCTGTTCTGTACAGAGAC	1490
Dp	1490	TGGGGGATGACACAAGCTGTGCTCACAGACACCTGTATTTCTGTTCTGTACAGAGAC	1549
QY	1491	CAGCTTGCCCTTTAAGATGTGCAAAACATGTGCCCTGATCTCGTTGGGCTGTGGCCATAT	1550
Dp	1550	CAGCTTGCCCTTTAAGATGTGCAAAACATGTGCCCTGATCTCGTTGGGCTGTGGCCATAT	1609
QY	1551	TTGTCACTGTATCTCCCTCTTGGTGTCAAAAAACAACAAGATATCAACCCAAATAGAA	1610
Dp	1610	TTGTCACTGTATCTCCCTCTTGGTGTCAAAAAACAACAAGATATCAACCCAAATAGAA	1669
QY	1611	ATAGTCTGTGGAAATGTGTGACAAAGCAAGGCTGTAGTGTCTTCTCAACCGTGCAAAAG	1670
Dp	1670	ATAGTCTGTGGAAATGTGTGACAAAGCAAGGCTGTAGTGTCTTCTCAACCGTGCAAAAG	1729
QY	1671	CCGTGTCTTCCCGGGAAACCAAGAAAGATCCGTACTCAAAAACCAAGATTTAAAG	1730
Dp	1730	CCGTGTCTTCCCGGGAAACCAAGAAAGATCCGTACTCAAAAACCAAGATTTAAAG	1789
QY	1731	GAGTTCCTTAAATTTGCACTTGTGTTCTGGAAGCTACTTTTCAGTGCCATTTGATGTGGA	1790
Dp	1790	GAGTTCCTTAAATTTGCACTTGTGTTCTGGAAGCTACTTTTCAGTGCCATTTGATGTGGA	1849

QY	1791	TGTCGTGAGAGGCGATTAACCTTTTCCCTAAAGATATTTGGTAATGAGATATTTGG	1850
Db	1850	TGTCGTGAGATGGCTATTAACCTTTTTCCTAAAGATATTTGGTAATGATATTTGG	1909
QY	1851	TTTGGGGAAGTTGAAATTTTATAGGTTAAATGTCATTTAGAGATGGGAGAGATTA	1910
Db	1910	TTTGGGGAAGTTGAAATTTTATAGGTTAAATGTCATTTAGAGATGGGAGAGAGATTA	1969
QY	1911	TACTGCAGGAGCTTCAGCCATGTTGGAAACTGATAAAGCACTTAGCAAGGCTTCT	1970
Db	1970	TACTGCAGGAGAGCTTCAGCCATGTTGGAAACTGATAAAGCACTTAGCAAGGCTTCT	2029
QY	1971	TTCAATTTTATTTATGTTTCACTATATAAAGCTTAGGAACTAGTAGGATGAACAAC	2030
Db	2030	TTCAATTTTATTTATGTTTCACTATATAAAGCTTAGGAACTAGTAGGATGAACAAC	2088
QY	2031	TGTCGCCAGAGTAAGAGAGAGAGCTACTATATGATTTAGAGCCATAACCGATTAC	2090
Db	2090	TGTCGCCAGAGTAAGAGAGAGAGCTACTATATGATTTAGAGCCATAACCGATTAC	2143
QY	2091	GAAAGAGGCGGATTACTTTCAGCTTCCATGTAACGTATGCAATTAAGCAATGTA	2150
Db	2150	GAAAGAGGCGGATTACTTTCAGCTTCCATGTAACGTATGCAATTAAGCAATGTA	2209
QY	2151	GTTTCTAAGATCAAGTTCGCAAGCTAACGTAATCCCATCTCAATCACACTATGA	2210
Db	2210	GTTTCTAAGATCAAGTTCGCAAGCTAACGTAATCCCATCTCAATCACACTATGA	2266
QY	2211	TGATGGAACAATAACAGGCCCAAGCCTGTGGTATGATGTGCACATTTGCTAG	2270
Db	2270	TGATGGAACAATAACAGGCCCAAGCCTGTGGTATGATGTGCACATTTGCTAG	2322
QY	2271	AAAATPACTCTCTCATTAATGGGTTGGAGATATTTGGTGAACAACCTACTTGG	2330
Db	2330	AAAATPACTCTCTCATTAATGGGTTGGAGATATTTGGTGAACAACCTACTTGG	2386
QY	2331	AGTAGAAGGAATGATTAATTCATTAATTCATTAATTCATGAGACATTTAGTAG	2399
Db	2390	AGTAGAAGGAATGATTAATTAATTAATTAATTCATTAATTCATGAGACATTTAG	2444
QY	2391	TATACCAAGCGATGATGCTGAGTCACTCTGTGTATATTTTCCAAATTTTGTAT	2450
Db	2450	TATACCAAGCGATGATGCTGAGTCACTCTGTGTATATTTTCCAAATTTTGTAT	2509
QY	2451	CTGCACATATTTGAAATCAAAATATTAAGACTTTCGAAAATTTGGTCCCTGG	2510
Db	2510	CTGCACATATTTGAAATC-ATAATATTAAGACTTTCGAAAAGATGAGGTCCTGG	2566
QY	2511	TGGCAACTTGTACGTAGAGATTTCCCTGGTTGGAACTTAACCATTTACATAT	2570
Db	2569	TGGCAACTTGTACGTAGAGATTTCACTCTGTGTGTAGATCTTAACCATTTACAT	2622
QY	2571	TAGACAAGACATTTTATTTTCTTCTCTGTAAGAAA-AAAATGAGGGAAGAGACA	2624
Db	2629	TAGACATGACATTTCTTTCTCTCTCTCTGTAAGAAAATTAAGTGGGAAGAGACA	2663
RESULT 12			
US-10-232-224-41			
; Sequence 41, Application US/10232224			
; Publication No. US20030065147A1			
; GENERAL INFORMATION:			
; APPLICANT: Baker, Kevin P.			
; APPLICANT: Desnoyers, Luc			
; APPLICANT: Geritsen, Mary			
; APPLICANT: Goddard, Audrey			
; APPLICANT: Godowski, Paul J.			
; APPLICANT: Grimaldi, J. Christopher			
; APPLICANT: Guiney, Austin L.			
; APPLICANT: Smith, Victoria			
; APPLICANT: Stephan, Jean-Philippe F.			
; APPLICANT: Matanabe, Colin L.			
; APPLICANT: Wood, William I.			

QY	568	GATGGAATTTCAATCTAAGCTCTTCCACACACTTGGTCAGATATTTCCAGAAATTGGAGCGAT	627
Db	590	GATGGAATTTCAATCTAAGCTCTTCCACACACTTGGTCAGATATTTCCAGAAATTGGAGCGAT	649
QY	628	GTTCAAGTGAAGATCTGTGGACACAGCCCAATGTGCACTTGGGCGCTCAACTCATGGAAG	687
Db	650	GTTCAAGTGAAGATCTGTGGACACAGCCCAATGTGCACTTGGGCGCTCAACTCATGGAAG	709
QY	688	TGACTGTCTACAGAAACATGSAACGGCATATGTTCCATGSCACAAAGTGAAGATGTGT	747
Db	710	TGACTGTCTACAGAAACATGSAACGGCATATGTTCCATGSCACAAAGTGAAGATGTGT	769
QY	748	ACGTGTAAACAGATCAGATTCCTGAGTTTGTACTATGTTCCAGAAAGCATTCGAAAT	807
Db	770	ACGTGTAAACAGATCAGATTCCTGAGTTTGTACTATGTTCCAGAAAGCATTCGAAAT	829
QY	808	CATCCGACGAACCTTCC - CAAAGATCTCCCATATGTTGATGTCTGATTCATGATC	866
Db	830	CATCCGACGAACCTTCCAAAGATCTCCCATATGTTGATGTCTGATTCATGATC	889
QY	867	CTAGCAGCTTCTCATATATTTCTACATTAACATCAAGGAGCTTCGGGATTAATCTAGC	926
Db	890	CTAGCAGCTTCTCATATATTTCTACATTAACATCAAGGAGCTTCGGGATTAATCTAGC	949
QY	927	GCGTGTGTGTTCCACCAATCATATCTGGAATCACAGTATGTGCTCAATGSAACCTTCA	986
Db	950	GCGTGTGTGTTCCACCAATCATATCTGGAATCACAGTATGTGCTCAATGSAACCTTCA	1008
QY	987	GCGTTAACTCACTGTGAAGGTGAGACACAGAGCACTGTCCGCAACGACACCAACAC	1044
Db	1010	GCGTTAACTCACTGTGAAGGTGAGACACAGAGCACTGTCCGCAACGACACCAACAC	1066

DB 1070 CCAGACCTTCAAAAACCCACCCCTCTTTAGCAACTACTCTAAAACTATGATTCAACA 112

QY	1075	-----AGGACCTGCTGGTACCAACCCCTGGAGCTGAGTAGATTCCTATGTAAAACTCGC	1113
Dp	1130	CCCCGAGACCTAATGGTGTACAAACCCCTGGAGCTGTAGATTCCTGATGTAAAACTGCG	1189
QY	1131	AGATTAAAGATATGGCCACTTTCAAGCCACATCAAAATTGTAGAGGAATCTTAAAGG	1197
Dp	1190	AGATTAAAGATATGGCCACTTTCAAGCCACATCAAAATTGTAGAGGAATCTTAAAGG	1244
QY	1191	TTAAATCATCATCAAGATGACAGAGCTCTGATGCCGTGCCAATGGCTTAAAACTCCCTAA	1250
Dp	1250	TTAAATCATCATCAAGATGACAGAGCTCTGATGCCGTGCCAATGGCTTAAAACTCCCTAA	1309
QY	1251	TGACCTTTGTGTGTCGTCCTGCCAAGGAGGACATTCCTCAAGAGAGTCTGTACCAATCTTCTG	1311
Dp	1310	TGACCTTTGTGTGTCGTCCTGCCAAGGAGGACATTCCTCAAGAGAGTCTGTACCAATCTTCTG	1366
QY	1311	ACCCCACTCTGGAGATCAACCCGAAACAAGTCTGCAGCCCTGTGGATGTGATGAGATGT	1370
Dp	1370	ACCCCACTCTGGAGATCAACCCGAAACAAGTCTGCAGCCCTGTGGATGTGATGAGATGT	1422
QY	1371	GTCGTGACCTGTGAGAGAACTTCAATGGGTGTGGAGCGTACTGTGAACCTCAACC	1430
Dp	1430	GTCGTGACCTGTGAGAGAACTTCAATGGGTGTGGAGCGTACTGTGTGAACCTCAACC	1488
QY	1431	TGGGGGATGACACAAGCCTGGCTCTCAAGACACCTGATTTCTGTTCTGTACAGAGACC	1490
Dp	1490	TGGGGGATGACACAAGCCTGGCTCTCAAGACACCTGATTTCTGTTCTGTACAGAGACC	1548
QY	1491	CAGCTGTGACCTTTAAGGATGTGCAAAAGTGCCCTGATCTCCGTTGGCTGCTTGGCCATAT	1551
Dp	1550	CAGCTGTGACCTTTAAGGATGTGCAAAAGTGCCCTGATCTCCGTTGGCTGCTTGGCCATAT	1609
QY	1551	TTGTCACTGTATCTCCCTCTTGTGTGTACAAAAACAAGAAATACAAACCAATATAGAA	1610
Dp	1610	TTGTCACTGTATCTCCCTCTTGTGTGTACAAAAACAAGAAATACAAACCAATATAGAA	1666

1670 ATAGCCCTGGAAATGTGTGCAAGAAAGGCTGAGTGTCTTCTCAACCGTGAAG 1729  
1671 CCGTGTCTTCCCGGAAACAGAAAGATCCCTACTCAAAAACAGATTTAAAG 1730  
1730 CCGTGTCTTCCCGGAAACAGAAAGATCCCTACTCAAAAACAGATTTAAAG 1789  
1731 GAGTTCTTAAATTTGACCTTTGTTGTAAGCTCAGTTTCAGTGCATTGATGAGA 1790  
1790 GAGTTCTTAAATTTGACCTTTGTTGTAAGCTCAGTTTCAGTGCATTGATGAGA 1849  
1791 TGTGTGAGATGAGTAACTTTTTCCTTAAGATTTATTTGTAATTAATTTGAG 1850  
1850 TGTGTGAGATGAGTAACTTTTTCCTTAAGATTTATTTGTAATTAATTTGAG 1909  
1851 TTGGGGAGTGAATTTTTTAAGTTAAATGATTTTGAATGAGATGGGAGAGGATTA 1910  
1910 TTGGGGAGTGAATTTTTTAAGTTAAATGATTTTGAATGAGATGGGAGAGGATTA 1969  
1911 TACTGAGGAGCTTCAGCCATGTTGTGAACTGATTAAGCACTTACAGGCTTCTT 1970  
1970 TACTGAGGAGCTTCAGCCATGTTGTGAACTGATTAAGCACTTACAGGCTTCTT 2029  
1971 TTCAATATTTTAAATGTTTCACTTTAAAGTCTTGTGTAAGTAAAGCACTG 2030  
2030 TTCAATATTTTAAATGTTTCACTTTAAAGTCTTGTGTAAGTAAAGCACTG 2089  
2031 TGTCCGAGAGTAAGAGAGAGTACTGATTAAGAGCTTACCCAGGTTAACTGCA 2090  
2090 TGTCCGAGAGTAAGAGAGAGTACTGATTAAGAGCTTACCCAGGTTAACTGCA 2149  
2091 GAAGAGGCGGATCTTCACTTCACTGATTAAGTAAAGTAAAGCCATGATGACA 2150  
2150 GAAGAGGCGGATCTTCACTTCACTGATTAAGTAAAGTAAAGCCATGATGACA 2209  
2151 GTTTCTAAGATCATGTTCCAGTAACTGATTAAGTAAAGTAAAGCCATGATGACA 2210  
2210 GTTTCTAAGATCATGTTCCAGTAACTGATTAAGTAAAGTAAAGCCATGATGACA 2269  
2211 TGATGAACATTAACAGGCGGAGCTGTGATGATGAGACACTTGTCAACACAGAA 2270  
2270 TGATGAACATTAACAGGCGGAGCTGTGATGATGAGACACTTGTCAACACAGAA 2329  
2271 AAAAATCTACTCTCATTAATGGGTGGAGATTTTGGTGAACAACCTACTTGTGCTG 2330  
2330 AAAAATCTACTCTCATTAATGGGTGGAGATTTTGGTGAACAACCTACTTGTGCTG 2389  
2331 AGTGAAGATGATATTTCAATATTTATTTATTTCCATGACATTTAGTGTCTTTTA 2390  
2390 AGTGAAGATGATATTTCAATATTTATTTATTTCCATGACATTTAGTGTCTTTTA 2449  
2391 TATACAGGAGATGATCTGATGAGACCTCTGTGATATTTCCAAATTTTGTATAGTGA 2450  
2450 TATACAGGAGATGATCTGATGAGACCTCTGTGATATTTCCAAATTTTGTATAGTGA 2509  
2451 CTGACATATTTGAAATCAAAATATTAAGATTTTCCAAATTTTGTATAGTGA 2510  
2510 CTGACATATTTGAAATCAAAATATTAAGATTTTCCAAATTTTGTATAGTGA 2568  
2511 TGGCAACTTGATAGTAAGATTTCCCTCTGTTTGAAGTAAACCATTTACTATATGT 2570  
2569 TGGCAACTTGATAGTAAGATTTCCCTCTGTTTGAAGTAAACCATTTACTATATGT 2628  
2571 TGAGCAAGACATTTTCTTCTCTCTGAAATAA-AAAAATGAGGAGAGAGACA 2624  
2629 TAGACATGACATTTCTTCTCTCTCTGAAATAAATTAAGTGTGGAGAGAGACA 2683

RESULT 13  
US-10-227-884-41  
; Sequence 41, Application US/10227884  
; Publication No. US20030027988A1  
; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Desnoyers, Luc  
APPLICANT: Gerritsen, Mary  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Smith, Victoria  
APPLICANT: Stephan, Jean-Philippe F.  
APPLICANT: Watanabe, Colin L.  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
TITLE OF INVENTION: ACTS ENCODING THE SAME  
FILE REFERENCE: P3530P1C79  
CURRENT APPLICATION NUMBER: US/10/227, 884  
CURRENT FILING DATE: 2002-08-26  
PRIOR APPLICATION NUMBER: 10/119, 480  
PRIOR FILING DATE: 2002-04-09  
PRIOR APPLICATION NUMBER: 60/059113  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/062287  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/063549  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/064103  
PRIOR FILING DATE: 1997-10-31  
PRIOR APPLICATION NUMBER: 60/069873  
PRIOR FILING DATE: 1997-12-17  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/086392  
PRIOR FILING DATE: 1998-05-22  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089905  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/090472  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090557  
PRIOR FILING DATE: 1998-06-24  
PRIOR APPLICATION NUMBER: 60/090691  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/090695  
PRIOR FILING DATE: 1998-06-25  
PRIOR APPLICATION NUMBER: 60/091982  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/095302  
PRIOR FILING DATE: 1998-08-04  
PRIOR APPLICATION NUMBER: 60/095318  
PRIOR FILING DATE: 1998-08-04  
PRIOR APPLICATION NUMBER: 60/095916  
PRIOR FILING DATE: 1998-08-10  
PRIOR APPLICATION NUMBER: 60/096146



Db 290 AAACTCTGGAAAGGAGCCGTGTGACAGCGGTCTGTACCAAGTACTCACACGCCCTCG 349  
Qy 328 TGGGCTCAAAATATTAACATTTGGCGGTGAACCTGATTTTCCCTAGATGGCAAAAGAAATG 387  
Db 350 TGGGTCTCAAAATATAACATTTGGCGGTGAACCTGATTTTCCCTAGATGGCAAAAGAAATG 409  
Qy 368 CCAATGGGAACATAGTCTATGAGAGAACTGCAAGAAATGAGCTGGTTATCTGTATC 447  
Db 410 CCAATGGCAACATAGTCTATGAGAGAACTGCAAGAAATGAGCTGGTTATCTGTATC 469  
Qy 448 CATATGTTTACAAGTGAACAGCATGTCAGAGACAGTGAACGGGAAAAATGGCACCGGCC 507  
Db 470 CGTATGTTTACAAGTGAACAGCATGTCAGAGACAGTGAACGGGAAAAATGGCACCGGCC 529  
Qy 508 AAAGCCATCATACGTCTTCCCTGATGGAAAACCTTTTCTTCCACACCCCGGATGGAGAA 567  
Db 530 AAAGCCATCATACGTCTTCCCTGATGGAAAACCTTTTCTTCCACACCCCGGATGGAGAA 589  
Qy 568 GATGGAAATTCATCTACGTCTTCCCAACAAGTGGTCACTTGGTCACTTCCAGAAATGGAGCAT 627  
Db 590 GATGGAAATTCATCTACGTCTTCCCAACAAGTGGTCACTTGGTCACTTCCAGAAATGGAGCAT 649  
Qy 628 GTTCACTGAGAGTTTCTGTGAACAAGCCCAATGTGACACTTGGGCTCACTCATGGAAG 687  
Db 650 GTTCACTGAGAGTTTCTGTGAACAAGCCCAATGTGACACTTGGGCTCACTCATGGAAG 709  
Qy 688 TCACTGTCTACAGAGACATGGAACGGGCAATGTTCCCATGCGCAACAAGTGAAGATGTGT 747  
Db 710 TCACTGTCTACAGAGACATGGAACGGGCAATGTTCCCATGCGCAACAAGTGAAGATGTGT 769  
Qy 748 AGCTGGTACAGATCGATTCCTGTGTGTGTGACATATGTTCCAGAGAAAGCATGCAAT 807  
Db 770 AGCTGGTACAGATCGATTCCTGTGTGTGTGACATATGTTCCAGAGAAAGCATGCAAT 829  
Qy 808 CATCCAGCAAAACCTTCC-CAAAGATCTCCCATATATGTTGATGTCTGATTCATGATC 866  
Db 830 CATCCAGCAAAACCTTCCCAAGATCTCCCATATATGTTGATGTCTGATTCATGATC 889  
Qy 867 CTAGCCACTTCTCTCAATATTTCTACCACTTACTACAAAGTGAAGCTTGGGGAAATATATG 926  
Db 890 CTAGCCACTTCTCTCAATATTTCTACCACTTACTACAAAGTGAAGCTTGGGGAAATATATG 949  
Qy 927 GCTGTGTTGTTTCCACCAATCACTACTGTGTAATCAACGTAATGTCTCAATGTAACCTTCA 986  
Db 950 GCTGTGTTGTTTCCACCAATCACTACTGTGTAATCAACGTAATGTCTCAATGTAACCTTCA 1009  
Qy 987 GCTGTAACTCACTGTGAAGAGTGCAGACACAGACCTTGTCCGCCACGGCCACACAC 1046  
Db 1010 GCTGTAACTCACTGTGAAGAGTGCAGACACAGACCTTGTCCGCCACGGCCACACAC 1069  
Qy 1047 CCAAGCTTCAAAACCAACCCCTTCTT----- 1074  
Db 1070 CCAAGCTTCAAAACCAACCCCTTCTTCTTAACTACTCTAAATTTTATGATCAAAACA 1129  
Qy 1075 -----AGAGCCCTGGGAGCAACCCCTGGAGCTGATAGATTCCTGATGAAAACTGCC 1130  
Db 1130 CCCCAGAGCTTACTGTGTGAACACCCCTGGAGCTGATAGATTCCTGATGAAAACTGCC 1189  
Qy 1131 AGATTAACAGATATGAGCACTTTCAAGCCACCATCAATCAATTTGAGAGGAATCTTTAAGG 1190  
Db 1190 AGATTAACAGATATGAGCACTTTCAAGCCACCATCAATTTGAGAGGAATCTTTAAGG 1249  
Qy 1191 TTAACATCATCAAGATGACAGAGCTGCTGATGACGGGTGCCATGGCTGAAAGCTCCCTAA 1250  
Db 1250 TTAACATCATCAAGATGACAGAGCTGCTGATGACGGGTGCCATGGCTGAAAGCTCCCTAA 1309  
Qy 1251 TAGACTTTGTGTGCTGCGGCAAGAGGAGCATTTCCACGAGAGTCTGTACCATATTTCTG 1310  
Db 1310 TAGACTTTGTGTGCTGCGGCAAGAGGAGCATTTCCACGAGAGTCTGTACCATATTTCTG 1369  
Qy 1311 ACCCAGCTGCGAGATCAACCAAGCAACAGTCTGAGCTTGTGATGTGATGATGT 1370

Db 1370 ACCCAGCTGCGAGATCAACCAAGCAACAGTCTGAGACCCCTGTGATGTGATGATGT 1429  
Qy 1371 GTCTGTGACTGTGACAGCAACCTTCAATGAGTGTGGAGACGTACTGTGTGAACCTCACCC 1430  
Db 1430 GTCTGTGACTGTGACAGCAACCTTCAATGAGTGTGGAGACGTACTGTGTGAACCTCACCC 1489  
Qy 1431 TGGGGGATGACACAGAGCTGGCTCTCAAGAGCAACCTGATTTCTGTTCTGACAGAGACC 1490  
Db 1490 TGGGGGATGACACAGAGCTGGCTCTCAAGAGCAACCTGATTTCTGTTCTGACAGAGACC 1549  
Qy 1491 CAGCCTGCTTTAAGATGGCAAAAGTGGCTGTATCTCGTTGGCTCTTGGCCATAT 1550  
Db 1550 CAGCCTGCTTTAAGATGGCAAAAGTGGCTGTATCTCGTTGGCTCTTGGCCATAT 1609  
Qy 1551 TTGTCACTGTATCTCCTCTTGGTGAACAAAAACAAGAAATACAAACCAATATGAAA 1610  
Db 1610 TTGTCACTGTATCTCCTCTTGGTGAACAAAAACAAGAAATACAAACCAATATGAAA 1669  
Qy 1611 ATAGTCTGGGAATGCTGTGACAAAGGCTGAGTGTCTTCTCAACGCTGCANAAG 1670  
Db 1670 ATAGTCTGGGAATGCTGTGACAAAGGCTGAGTGTCTTCTCAACGCTGCANAAG 1729  
Qy 1671 CCGTGTCTTCCCGGAAACAGAAAGATCCGCTACTCAAAAAACCAAGAAATTTAAG 1730  
Db 1730 CCGTGTCTTCCCGGAAACAGAAAGATCCGCTACTCAAAAAACCAAGAAATTTAAG 1789  
Qy 1731 GAGTTCTTAAATTTGCACTTGTGTTCTGAGTCACTTTTCAAGGCCATTTGATGGA 1790  
Db 1790 GAGTTCTTAAATTTGCACTTGTGTTCTGAGTCACTTTTCAAGGCCATTTGATGGA 1849  
Qy 1791 TGTGTGAGTGTGATTAACCTTTTCTTAAAGATTAATGTTAAATAGATTAATGTGTG 1850  
Db 1850 TGTGTGAGTGTGATTAACCTTTTCTTAAAGATTAATGTTAAATAGATTAATGTGTG 1909  
Qy 1851 TTTGGGGAAGTGAATTTTCTTAAAGTCAATTTTAAAGATGAGGAGGAGTTA 1910  
Db 1910 TTTGGGGAAGTGAATTTTCTTAAAGTCAATTTTAAAGTCAATTTTAAAGATGAGGAGTTA 1969  
Qy 1911 TACTGAGGAGCTTCAAGCATGTTGTGAACCTGATTAAGCACTTAAAGGCTTCTT 1970  
Db 1970 TACTGAGGAGCTTCAAGCATGTTGTGAACCTGATTAAGCACTTAAAGGCTTCTT 2029  
Qy 1971 TTTCAATTTTCTTAACTTCACTTAAAGTCTTAAAGTCTTAAAGTCTTAAAGTCTTAA 2030  
Db 2030 TTTCAATTTTCTTAACTTCACTTAAAGTCTTAAAGTCTTAAAGTCTTAAAGTCTTAA 2089  
Qy 2031 TGTCCGAGAGTGAAGAGAGAGTACTATTTGATTAAGGCTTAAACCGGTTAACTGCAA 2090  
Db 2090 TGTCCGAGAGTGAAGAGAGAGTACTATTTGATTAAGGCTTAAACCGGTTAACTGCAA 2149  
Qy 2091 GAAGAGCGGGATACTTTCAGCTTTCATGTAACCTGATTAAGGCTTAAACCGGTTAACTGCAA 2150  
Db 2150 GAAGAGCGGGATACTTTCAGCTTTCATGTAACCTGATTAAGGCTTAAACCGGTTAACTGCAA 2209  
Qy 2151 GTTTCTTAAGTCAATGTTCCAGCTTCACTGATTAACCTTCAATCACTCACTCACTGCTCC 2210  
Db 2210 GTTTCTTAAGTCAATGTTCCAGCTTCACTGATTAACCTTCAATCACTCACTCACTGCTCC 2269  
Qy 2211 TGAATGAACATTAACAGGCCCAAGCTGTGTGATGATGTGACACTTGTGACTGAGAA 2270  
Db 2270 TGAATGAACATTAACAGGCCCAAGCTGTGTGATGATGTGACACTTGTGACTGAGAA 2329  
Qy 2271 AAAATACTACTCTCATTAATGAGGTGGAGATTTTGTGTGACAACTTGTGCTGGCTG 2330  
Db 2330 AAAATACTACTCTCATTAATGAGGTGGAGATTTTGTGTGACAACTTGTGCTGGCTG 2389  
Qy 2331 AGTGAAGATGATTAATCATTTATTTATTTTCAAGCACTTAACTTAACTTAACTTAACTT 2390  
Db 2390 AGTGAAGATGATTAATCATTTATTTATTTTCAAGCACTTAACTTAACTTAACTTAACTT 2449  
Qy 2391 TATACAGAGCATGATGTGATGACACTCTTGTGTATTTTCAAAATTTTGTATATGTG 2450  
Db 2450 TATACAGAGCATGATGTGATGACACTCTTGTGTATTTTCAAAATTTTGTATATGTG 2509









QY 1075 ----AGGACCTGCTGTGACAAACCCCTGGAGCTGAGTAGGATTCTGATGAAAACCTGCC 1130  
 Db 1130 CCCAGAGACCTACTGAGTGAACAACCCCTGAGAGCTGAGTAGGATTCTGATGAAAACCTGCC 1189  
 QY 1131 AGATTAACAGATATGCGCACTTTCAAGCCACCATCACATTGTAGAGGATCTTAGAGG 1190  
 Db 1190 AGATTAACAGATATGCGCACTTTCAAGCCACCATCACATTGTAGAGGATCTTAGAGG 1249  
 QY 1191 TTATCATCATCATGACAGACGCTGATGCGGAGGCTTGGCCCTGAAAGCTCCCTAA 1250  
 Db 1250 TTATCATCATCATGACAGACGCTGATGCGGAGGCTTGGCCCTGAAAGCTCCCTAA 1309  
 QY 1251 TAGACTTTGTGTGACCTGGCAAGGAGGACATTCCACGAGAGTGTGATACCATCTTTCTG 1310  
 Db 1310 TAGACTTTGTGTGACCTGGCAAGGAGGACATTCCACGAGAGTGTGATACCATCTTTCTG 1369  
 QY 1311 ACCCCACCTGCGAGATACCCGAGAACAGCTGTGACAGCCCTGTGATGTGATGATGT 1370  
 Db 1370 ACCCCACCTGCGAGATACCCGAGAACAGCTGTGACAGCCCTGTGATGTGATGATGT 1429  
 QY 1371 GTCTGCTGACGTGAGAGCAACCTTCATGGGTCTGGGAGGTACTGTGTGAACCTCAACC 1430  
 Db 1430 GTCTGCTGACGTGAGAGCAACCTTCATGGGTCTGGGAGGTACTGTGTGAACCTCAACC 1489  
 QY 1431 TGGGGGATGACACAAAGCCTGGCTCTACAGACACCCCTGATTTCTGTTCTGACAGAGACC 1490  
 Db 1490 TGGGGGATGACACAAAGCCTGGCTCTACAGAGACCCCTGATTTCTGTTCTGACAGAGACC 1549  
 QY 1491 CAGCCTGCTTTAAGATGTGCAAAAGTGTGCTGATCTCGTGTGCTGCTGCTGCTAT 1550  
 Db 1550 CAGCCTGCTTTAAGATGTGCAAAAGTGTGCTGATCTCGTGTGCTGCTGCTGCTAT 1609  
 QY 1551 TTGTCACTGATCTCCCTCTGTGTGACAAAACCAAGAAATCAACCCATAGAAA 1610  
 Db 1610 TTGTCACTGATCTCCCTCTGTGTGACAAAACCAAGAAATCAACCCATAGAAA 1669  
 QY 1611 ATAGCTCGGGAATGTGTGAGAGCAAAAGCCTGAGTGTCTTCTCAACCGTGCAAAAG 1670  
 Db 1670 ATAGCTCGGGAATGTGTGAGAGCAAAAGCCTGAGTGTCTTCTCAACCGTGCAAAAG 1729  
 QY 1671 CCGTGTCTTCCCGGGAACCAAGAAAGATCCGCTACTCAAAAACCAAGAAATTTAAAG 1730  
 Db 1730 CCGTGTCTTCCCGGGAACCAAGAAAGATCCGCTACTCAAAAACCAAGAAATTTAAAG 1789  
 QY 1731 GAGTTCTTAATTTGAGACCTGTCTGAAAGCTCACTTTTCAGTCCATGATGTGAGA 1790  
 Db 1790 GAGTTCTTAATTTGAGACCTGTCTGAAAGCTCACTTTTCAGTCCATGATGTGAGA 1849  
 QY 1791 TGTGCTGAGGTGCTATTAACCTTTTTCCTAAAGATTAATGTAAATAGATATTGTGG 1850  
 Db 1850 TGTGCTGAGGTGCTATTAACCTTTTTCCTAAAGATTAATGTAAATAGATATTGTGG 1909  
 QY 1851 TTTGGGGAAGTGAATTTTATAGGTGAATGCTATTTTGAAGATGGGAGGAGATTA 1910  
 Db 1910 TTTGGGGAAGTGAATTTTATAGGTGAATGCTATTTTGAAGATGGGAGGAGATTA 1969  
 QY 1911 TACTGAGGAGCTTACGCAATGTGTGAACCTGAATAAAGCAACTTACCAAGGCTTCTT 1970  
 Db 1970 TACTGAGGAGCTTACGCAATGTGTGAACCTGAATAAAGCAACTTACCAAGGCTTCTT 2029  
 QY 1971 TTCAATTAATTTTATGTTTCACTTAATTAAGCTTTAGTAAGTAGTAGTAAGAACACG 2030  
 Db 2030 TTCAATTAATTTTATGTTTCACTTAATTAAGCTTTAGTAAGTAGTAGTAAGAACACG 2089  
 QY 2031 TGTCCCGAGATGAAGAGAAAGCTACTTGTATTAAGACCTTACCAGGTTTACTGCA 2090  
 Db 2090 TGTCCCGAGATGAAGAGAAAGCTACTTGTATTAAGACCTTACCAGGTTTACTGCA 2149  
 QY 2091 GAAGAGGCGGATACCTTACGCTTCCATGATAGTATGATGAAGCCAAAGTGTGCTCA 2150  
 Db 2150 GAAGAGGCGGATACCTTACGCTTCCATGATAGTATGATGAAGCCAAAGTGTGCTCA 2209

QY 2151 GTTCTTAAGATCAATGTTTCCAAAGCTAATGCAATGCCACTTCAATACACATCTGAACTCC 2210  
 Db 2210 GTTCTTAAGATCAATGTTTCCAAAGCTAATGCAATGCCACTTCAATACACATCTGAACTCC 2269  
 QY 2211 TGATGAACATATAACAGGCCCAAGCCTGTGTATGATGTGACACCTGTGATGACAA 2270  
 Db 2270 TGATGAACATATAACAGGCCCAAGCCTGTGTATGATGTGACACCTGTGATGACAA 2329  
 QY 2271 AAAATACTCTTCATTAATAGGTGTGAGATTTTGTGTGACAACTTCTGTGTGCTG 2330  
 Db 2330 AAAATACTCTTCATTAATAGGTGTGAGATTTTGTGTGACAACTTCTGTGTGCTG 2389  
 QY 2331 AGTGAAGGATGATTAATCATATATTCATTTATTCATGACATTTAGTATGCTTTTTA 2390  
 Db 2390 AGTGAAGGATGATTAATCATATATTCATTTATTCATGACATTTAGTATGCTTTTTA 2449  
 QY 2391 TATACGAGCATATGCTGATGACACACTTGTGTATATTTCCAAATTTTGTATGCTG 2450  
 Db 2450 TATACGAGCATATGCTGATGACACACTTGTGTATATTTCCAAATTTTGTATGCTG 2509  
 QY 2451 CTGCAATATTTGAATCAAAATATTAAGACTTTCCAAAAATTTGTCTGTTTCA 2510  
 Db 2510 CTGCAATATTTGAATCAAAATATTAAGACTTTCCAAAAATTTGTCTGTTTCA 2568  
 QY 2511 TGGCAACTGATGATGAAGGATTTCCCTGCTGTTTGAATTAACCATTTACTATATGT 2570  
 Db 2569 TGGCAACTGATGATGAAGGATTTCCCTGCTGTTTGAATTAACCATTTACTATATGT 2628  
 QY 2571 TAGACAAGACATTTTCTTCTCTGAAAAA-AAAATAGGAGAAAGAGACA 2624  
 Db 2629 TAGACAAGACATTTTCTTCTCTGAAAAA-AAAATAGGAGAAAGAGACA 2683

RESULT 15  
 US-10-230-338-41  
 ; Sequence 41, Application us/10230338  
 ; Publication No. US20030044934A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Gerritsen, Mary  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Grimaldi, J. Christopher  
 ; APPLICANT: Guiney, Aubert L.  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Stephan, Jean-Philippe F.  
 ; APPLICANT: Watanabe, Colin L.  
 ; APPLICANT: Wood, William I.  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; FILE OF INVENTION: ACIDS ENCODING THE SAME  
 ; FILE REFERENCE: P3550P1C92  
 ; CURRENT APPLICATION NUMBER: US/10/230,338  
 ; CURRENT FILING DATE: 2002-08-28  
 ; PRIOR APPLICATION NUMBER: 10/119,480  
 ; PRIOR FILING DATE: 2002-04-09  
 ; PRIOR APPLICATION NUMBER: 60/059113  
 ; PRIOR FILING DATE: 1997-09-17  
 ; PRIOR APPLICATION NUMBER: 60/062287  
 ; PRIOR FILING DATE: 1997-10-17  
 ; PRIOR APPLICATION NUMBER: 60/063549  
 ; PRIOR FILING DATE: 1997-10-28  
 ; PRIOR APPLICATION NUMBER: 60/064103  
 ; PRIOR FILING DATE: 1997-10-31  
 ; PRIOR APPLICATION NUMBER: 60/069873  
 ; PRIOR FILING DATE: 1997-12-17  
 ; PRIOR APPLICATION NUMBER: 60/078910  
 ; PRIOR FILING DATE: 1998-03-20  
 ; PRIOR APPLICATION NUMBER: 60/079294  
 ; PRIOR FILING DATE: 1998-03-25  
 ; PRIOR APPLICATION NUMBER: 60/079656  
 ; PRIOR FILING DATE: 1998-03-26  
 ; PRIOR APPLICATION NUMBER: 60/079728

```

: PRIOR FILING DATE: 1998-03-27
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 246
: SEQ ID NO 41
: LENGTH: 2683
: TYPE: DNA
: ORGANISM: Homo Sapien
US-10-230-338-41

```

Query Match	2485;	Score	2485;	DB	15;	Length	2683;
Best Local Similarity	97.8%;	Pred.	No. 0;				
Matches	2576;	Conservative	0;	Mismatches	20;	Indels	39;
						Gaps	4

QY	28	AACTGGTGGCTGGTCCGGAGAAATTCAGATGGAATGTCTCAATATTTCCGGAT	87
Db	50	AACCTGAATGGCTGGTCCGGAGAAATTCAGATGGAATGTCTCAATATTTCCGGAT	109
QY	88	TTTGTCTCTGGCTGCAGATTGCCACCTTGATGCCGCCAANCGATTTTCATGATGTGTGG	147
Db	110	TTTGTCTCTGGCTGCAGATTGCCACCTTGATGCCGCCAANCGATTTTCATGATGTGTGG	169
QY	148	GCAATGAAGAACCCTGTGCTTACATGAGGGAGACCAATCAATTAATGGCTGGTCTTCTG	207
Db	170	GCAATGAAGAACCCTGTGCTTACATGAGGGAGACCAATCAATTAATGGCTGGTCTTCTG	229
QY	208	ATGAATAATGACTGGAATGAAAACTCTACCCAGTGTGGAGCGGGAGACATGAGGTGGA	267
Db	230	ATGAATAATGACTGGAATGAAAACTCTACCCAGTGTGGAGCGGGAGACATGAGGTGGA	289
QY	268	AAAACTCTGGAAGGAGGCGGTGTGACAGGGGGTCTCTGACAGTGACTACACGACCTTG	327
Db	290	AAAACTCTGGAAGGAGGCGGTGTGACAGGGGGTCTCTGACAGTGACTACACGACCTTG	349
QY	328	TGGGCTCAAAATATACTTTGGCGGTGAACCTGTATAATCCCTGATGTCCAAAAGGAATG	387
Db	350	TGGGCTCAAAATATACTTTGGCGGTGAACCTGTATAATCCCTGATGTCCAAAAGGAATG	409
QY	388	CCAAATGGCAACATAGTCTATGAGMAGACTGCAAGAAATGAGGTGGTATATCTGCTGATC	447
Db	410	CCAAATGGCAACATAGTCTATGAGMAGACTGCAAGAAATGAGGTGGTATATCTGCTGATC	469
QY	448	CATATGTTTCAACACTGGAACAGATGGTTCAGAGACAGTACGGGGAAAAATGGCACCGGCC	507
Db	470	CATATGTTTCAACACTGGAACAGATGGTTCAGAGACAGTACGGGGAAAAATGGCACCGGCC	529
QY	508	AAAGCCATCATACGTTCTCCCTGATGGGAAACCTTTTCTCTACACCCCGAGTGAAGAA	567
Db	530	AAAGCCATCATACGTTCTCCCTGATGGGAAACCTTTTCTCTACACCCCGAGTGAAGAA	589
QY	568	GATGGAAATTTCAATCGTCTTCCACACACTGGTGAATTTCCAGAAATTGGGACGAT	627
Db	590	GATGGAAATTTCAATCGTCTTCCACACACTGGTGAATTTCCAGAAATTGGGACGAT	649
QY	628	GTTTCAGTGAAGTTTCTGGAACACACAGCCAAATGTGACATTTGGGCTCAACTCATGGAAG	687
Db	650	GTTTCAGTGAAGTTTCTGGAACACACAGCCAAATGTGACATTTGGGCTCAACTCATGGAAG	709
QY	688	TGAATGTCTTACAGAAACATGAGCGGGCATATGTTCCTCATCGACAAAGTAAAGATGTCT	747
Db	710	TGAATGTCTTACAGAAACATGAGCGGGCATATGTTCCTCATCGACAAAGTAAAGATGTCT	769
QY	748	ACGTGTGTAACAGTACGATCTCGTGTGTTGTGACTANGTTCCAGAGAAACGATCGAAATT	807
Db	770	ACGTGTGTAACAGTACGATCTCGTGTGTTGTGACTANGTTCCAGAGAAACGATCGAAATT	829
QY	808	CATCCGACGAACCTTCC-CAAAGATCTCCCATATGTTTGAATGTCTGATTCATGATC	866
Db	830	CATCCGACGAACCTTCCCTCCAAAGATCTCCCATATGTTTGAATGTCTGATTCATGATC	889
QY	867	CTAGGCACATTCCTCAATTTATTTCAACATTAACATCAAGGAGGAGCTTGGGGATTAATCTG	926
Db	890	CTAGGCACATTCCTCAATTTATTTCAACATTAACATCAAGGAGGAGCTTGGGGATTAATCTG	949

QY	927	GCCTGTTGTTTTCACCAATTCATCTGTAATCAACAGATGTGTCATGTGAACCTTCA	986
Ds	950	GCCTGTTGTTTTCACCAATTCATCTGTAATCAACAGATGTGTCATGTGAACCTTCA	1009
QY	987	GCCTTAACCTCACTGTGAAGAGTGCAGCAACAGAACCTTGTCCGCCACCCGACACAC	1046
Ds	1010	GCCTTAACCTCACTGTGAAGAGTGCAGCAACAGAACCTTGTCCGCCACCCGACACAC	1069
QY	1047	CCAGACCTTCAAAACCCACCCCTTCTTT-----	1074
Ds	1070	CCAGACCTTCAAAACCCACCCCTTCTTTAGCACTACTCTMAAATCTTTGATTCAAACA	1129
QY	1075	----AGGACCTGCTGTGAGCAACCCCTGAGAGTGGATGATCTCTGTAAGAACTGCC	1130
Ds	1130	CCCCAGACCTTCTGTGAGCAACCCCTGAGAGTGGATGATCTCTGTAAGAACTGCC	1189
QY	1131	AGATTACAGATATGCGCACTTTCAAGCCACCATCAAAATTGTAGAGGAGATCTTAGAG	1190
Ds	1190	AGATTACAGATATGCGCACTTTCAAGCCACCATCAAAATTGTAGAGGAGATCTTAGAG	1249
QY	1191	TTAACATCATCCAGATGACAGACGTCTGTATSCCGGTGCCATGGCTGAAGAGCTCCCTAA	1250
Ds	1250	TTAACATCATCCAGATGACAGACGTCTGTATSCCGGTGCCATGGCTGAAGAGCTCCCTAA	1309
QY	1251	TAGACTTTGTGCGACCTGCGCAAGGGAGCATTCOCAGAGGCTGTACCATCATTTCTG	1310
Ds	1310	TAGACTTTGTGCGACCTGCGCAAGGGAGCATTCOCAGAGGCTGTGTACCATCATTTCTG	1369
QY	1311	ACCCCACTTCGAGATCACCCAGACACACACTGTGCAGCCCTGTGTAGTGTGATGAGATGT	1370
Ds	1370	ACCCCACTTCGAGATCACCCAGACACACACTGTGCAGCCCTGTGTAGTGTGATGAGATGT	1429
QY	1371	GTCTGCTGACTGTGAGACGAACCTTCATGSGGTCTGGAGAGTACTGTGTGAACCTCACCC	1430
Ds	1430	GTCTGCTGACTGTGAGACGAACCTTCATGSGGTCTGGAGAGTACTGTGTGAACCTCACCC	1489
QY	1431	TGGGGGATGACACAGCCTGCTCTCACAGACACCTTGATTTCTGTTCCGACAGAGACC	1490
Ds	1490	TGGGGGATGACACAGCCTGCTCTCACAGACACCTTGATTTCTGTTCCGACAGAGACC	1549
QY	1491	CAGCTCGCCCTTAAGATGGCAACAGTGCCTGATCTCCGTGAGCTGCTGGCCATAT	1550
Ds	1550	CAGCTCGCCCTTAAGATGGCAACAGTGCCTGATCTCCGTGAGCTGCTGGCCATAT	1609
QY	1551	TTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACAGAGATACACCCCATAGAAA	1610
Ds	1610	TTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACAGAGATACACCCCATAGAAA	1669
QY	1611	ATPAGTCTGGGAATGTGTGCAGAAAGCAAGGCTGATGTCTTCTCAACCGTGCAAAAG	1670
Ds	1670	ATPAGTCTGGGAATGTGTGTGAGAAAGGCTGATGTCTTCTCAACCGTGCAAAAG	1729
QY	1671	CGGTGCTCTCCCGGGAAAAACAGAGAAAGATCCGCTACTCAAAAAACAGAAATTTAAAG	1730
Ds	1730	CGGTGCTCTCTCCCGGGAAAAACAGAGAAAGATCCGCTACTCAAAAAACAGAAATTTAAAG	1789
QY	1731	GAGTTTCTTAATTTGCACTTGTTCGAAGCTCACTTTCAGTGCATGTGATGTGAGA	1790
Ds	1790	GAGTTTCTTAATTTGCACTTGTTCGAAGCTCACTTTCAGTGCATGTGATGTGAGA	1849
QY	1791	TGTGCTGAGATGGCTATTAAACCTTTTTTCTTAAGAATATTTGTTAAATATGATTTGGG	1850
Ds	1850	TGTGCTGAGATGGCTATTAAACCTTTTTTCTTAAGAATATTTGTTAAATATGATTTGGG	1909
QY	1851	TTTGGGGAGTGTGAATTTTTTATAGSTTAATATCATTTTGAAGTGGGGAGAGGATTA	1910
Ds	1910	TTTGGGGAGTGTGAATTTTTTATAGSTTAATATCATTTTGAAGTGGGGAGAGGATTA	1969
QY	1911	TACTGCAGGACGCTTCAGCCATGTGTGTAAACTGATTAAGCAAACTTAGAGGCTTCTT	1970
Ds	1970	TACTGCAGGACGCTTCAGCCATGTGTGTAAACTGATTAAGCAAACTTAGAGGCTTCTT	2029
QY	1971	TTTCATTAATTTTATATGTTTCACTTAATAAGTCTTAGAGTACTAGTAGATGAAGAACTG	2030

Db	2030	TTGATATATTTTATATGTTTCACTTATAAGCTTAGGTAAGTATAGAAACACTG	2089
QY	2031	TGTCCGAGAGTAAGAGAGAAAGTACTATTTGATTAGAGCTAACCGAGTTAACTGCA	2090
Db	2090	TGTCCGAGAGTAAGAGAGAAAGTACTATTTGATTAGAGCTTAAGCCAGGTTAACTGCA	2149
QY	2091	GAAAGAGCGGAGATACCTTCAGCTTCATGTAACTGTATGATTAAGCCAAATGTATCCA	2150
Db	2150	GAAAGAGCGGAGATACCTTCAGCTTCATGTAACTGTATGATTAAGCCAAATGTATCCA	2209
QY	2151	GTTTCTAGATCAATGTTCCAAAGTAACTGAATCCCACTTCAATACACTCATGAAGTCC	2210
Db	2210	GTTTCTAGATCAATGTTCCAAAGTAACTGAATCCCACTTCAATACACTCATGAAGTCC	2269
QY	2211	TGATGGAACATATAACAGAGCCCAAGCCTGTGTATGATGTGACACACTGTGACTGAGAA	2270
Db	2270	TGATGGAACATATAACAGAGCCCAAGCCTGTGTATGATGTGACACACTGTGACTGAGAA	2329
QY	2271	AAATATCTACTCTCATTAATGGGTGGAGATATTTGGTGAACAACCTACTTGGCTGGCTG	2330
Db	2330	AAATATCTACTCTCATTAATGGGTGGAGATATTTGGTGAACAACCTACTTGGCTGGCTG	2389
QY	2331	AGTGAAGGATGATATTCATATATTCATTTATTCATGACATTTAGTTAGTCTTTTAA	2390
Db	2390	AGTGAAGGATGATATTCATATATTCATTTATTCATGACATTTAGTTAGTCTTTTAA	2449
QY	2391	TATACGAGCATGATGCTGAGTACACTCTTGTATATTTGCAAAATTTTGTATAGTGG	2450
Db	2450	TATACGAGCATGATGCTGAGTACACTCTTGTATATTTGCAAAATTTTGTATAGTGG	2509
QY	2451	CTGCACATATTTGAAATCAAAATATTAAGAATTTCACAAAATTTGGTCCCTGGTTTCA	2510
Db	2510	CTGCACATATTTGAAATCAAAATATTAAGAATTTCACAAAATTTGGTCCCTGGTTTCA	2568
QY	2511	TGGCACTTGATGATGAAGAATTTCCCTGTGTTGGAAGTAAACCATTTACTATATGT	2570
Db	2569	TGGCACTTGATGATGAAGAATTTCCCTGTGTTGGAAGTAAACCATTTACTATATGT	2628
QY	2571	TAGACAAGACATTTTCTTCTCTGAAAAA-AAAATGAGGGAAGAGACA	2624
Db	2629	TAGACAAGACATTTCTTCTCTGAAAAAATAAGTGTGGGAAGAGACA	2683

Search completed: September 10, 2004, 15:32:23  
 Job time : 3924 secs